The Dissertation Committee for Richard Paul Good
certifies that this is the approved version of the following dissertation:

The Stability and Performance of the EWMA and Double-EWMA Run-to-Run Controllers with Metrology Delay

Committee:

S. Joe Qin, Supervisor

Thomas F. Edgar

Gyeong S. Hwang

Gang Yu

Michael L. Miller

The Stability and Performance of the EWMA and Double-EWMA Run-to-Run Controllers with Metrology Delay

by

Richard Paul Good, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2004

Dedicated to Kara,
to Mom and Dad,
and to Heidi.

Acknowledgments

To start with, I wish to thank my advisor, Dr. S. Joe Qin, for his patience, guidance, and friendship throughout my time at the University of Texas. I would also like to thank my fellow students, especially Kevin Chamness, Jin Wang, Peter He, Jürgen Hahn, Chris and Scott Harrison, Elaine Hale, Henry Potrykus, Weilu Lin, Greg Cherry, Sebastian Lextrait, and Brent Bregenzer for the immeasurable help, the hours of discussion, and, most importantly, the countless distractions. I would also like to thank my fellow employees at AMD's Fab25 in Austin and Fab30 in Dresden, Germany. Without the help of Chris Bode, Jan Räbiger, Matt Purdy, Tom Sonderman, and Ernest Adams, I would not have been able to finish this dissertation while continuing to work in semiconductor manufacturing. I also wish to thank my committee, especially Michael Miller, and several anonymous paper reviewers for their priceless feedback in writing this dissertation. Finally I would like to thank my wonderful wife, Kara.

The Stability and Performance of the EWMA and Double-EWMA Run-to-Run Controllers with Metrology Delay

Publication No. _____

Richard Paul Good, Ph.D.
The University of Texas at Austin, 2004

Supervisor: S. Joe Qin

Because of the ever-increasing demands on product quality, feedback control has become a necessary enabling component in the manufacture of modern semiconductor devices. The nature of semiconductor manufacturing is such that measurements of device quality characteristics are not available during the processing of the product. Measurements are not made until after the product is processed and necessary changes to tool setting can only be made to subsequent production runs. This control scheme, termed run-to-run control, has become the cornerstone of process control in the semiconductor manufacturing industry.

In addition to the ever-increasing demands on product quality, the semiconductor manufacturing industry continues to see stringent growth in throughput requirements. Because of the demands on production throughput, it is rarely possible to perform quality measurements on a batch of wafers before processing begins on the following batch of wafers. The delay between product manufacturing and

product metrology coupled with inaccurate process models can lead to process instabilities and deterioration in controller performance. This dissertation investigates the robust stability requirements of processes controlled with EWMA and double-EWMA run-to-run controllers with delays between processing and metrology. In addition, the effects of model mismatch and metrology delay on the closed-loop performance of the EWMA and double-EWMA run-to-run controllers are derived by extending the robust stability methodology. Finally, these robust performance requirements are used to find the optimal tuning parameters for the double-EWMA controller. These tuning parameters allow for the largest model uncertainty while guaranteeing a predetermined minimum closed-loop transient performance.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Quality Control in Semiconductor Manufacturing

One of the largest challenges currently facing the semiconductor manufacturing industry is meeting the increasingly stringent requirements on product quality while maintaining the demands on product throughput. As the smallest features on semiconductor devices have been driven below $0.10\mu m$, quality control has exceedingly become an enabling technology in the manufacturing of these devices. The origins of quality control in semiconductor manufacturing are in statistical process control (SPC). The role of SPC is to chart key quality characteristics of manufacturing steps to determine when a product quality specification has fallen outside of a predetermined range. Once a processing fault has been identified, the root cause of the error is determined and a production engineer is responsible for making the necessary changes to a processing tool's setpoints in order to return the product to its quality specifications [42].

A fundamental assumption in SPC is that a process is stationary with an uncorrelated and normal distribution around its targeted operating point. This is often not the case, however, as many processes in semiconductor manufacturing experience autocorrelation and/or slow time-varying drifts to product quality. These drifts can be due to changes in processing at upstream manufacturing steps, degradation

in tool performance over time, ambient effects (from temperature, humidity, changes in atmospheric pressure, etc.) or changes to quality specifications. To compensate for these drifts in product quality, SPC has been combined with feedback control. This utilization of SPC for process control and improvement has been studied by Box and Kramer [9], Van der Weil *et al* [67], Tucker *et al* [65], Box and Luceno [10], and Janakiram and Keats [38] among others. The application of SPC for feedback control in semiconductor manufacturing was first introduced by Sachs *et al* [54] and Sachs, Hu, and Ingolfsson [55].

1.1.1 Run-to-Run Control

Tool aging and disturbances in upstream unit operations can cause semiconductor manufacturing processes to drift from their quality targets. It is, therefore, necessary to make corrective adjustments to tool settings. Since quality controlled parameters in most semiconductor manufacturing processes cannot be measured *in situ*, changes in tool settings can only be made between process runs. These settings are usually setpoints for a system of single input-single output PID control loops on the process tool where there is little or no visibility to the real-time wafer processing. As it is only possible to make changes to tool settings between process runs, these controllers are termed run-to-run (or run-by-run) controllers.

All run-to-run controllers consist of two fundamental components: an observer and a control law. The function of the observer is to estimate the states of a process which are usually the estimated parameters of an assumed process model or the estimates of a process disturbance. Once the observer estimates the states,

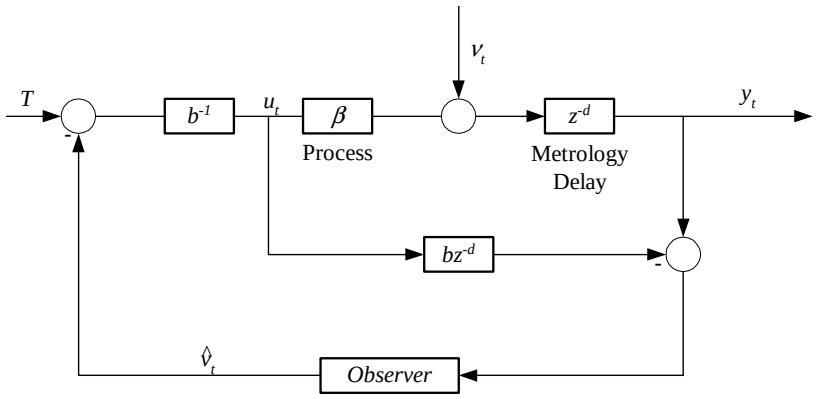

Figure 1.1: Schematic of a generic run-to-run controller

the control law calculates the inputs (or recipe) for the following batch of wafers. This control law is normally a simple inversion of an assumed process model so most run-to-run controllers are deadbeat controllers. A simple schematic of a run-to-run control system is shown in Figure 1.1.

1.1.1.1 The EWMA Controller

Because of its simplicity and robustness, the exponentially weighted moving average (EWMA) controller is the most common run-to-run controller found in semiconductor manufacturing. The EWMA controller was first applied to semiconductor manufacturing by Sachs, Hu, and Ingolfsson [55] to control a radially heated barrel epitaxy reactor. In this controller *ex situ* thickness uniformity measurements from a batch of wafers were used to update the recipe for the following batch of wafers. Since this pioneering work, run-to-run controllers have been successfully applied to many semiconductor manufacturing processes. Some of these

processes include chemical mechanical planarization [7, 12, 19], chemical deposition [59], plasma etching [34, 39], photolithography [6, 25], and rapid thermal processing [52].

Like all run-to-run controllers, the EWMA controller is composed of an observer and a control law. The EWMA observer assumes that the process output, y, is the linear combination of the process inputs, u, and a disturbance, ν,

$$y_t = \beta u_t + \nu_t \tag{1.1}$$

where β is the gain between the manipulated input and the measured output [55]. It is also assumed that the disturbance can be modeled as an integrated moving average (IMA{1,1}) time series of the form,

$$\nu_t = \nu_{t-1} - \theta\varepsilon_{t-1} + \varepsilon_t, \tag{1.2}$$

where ε is uncorrelated noise with variance σ_0^2. The function of the EWMA filter is to estimate the process disturbance from process data.

Box and Jenkins showed in [8] that the minimum mean squared error (MMSE) one step ahead estimate of an IMA{1,1} process is

$$\hat{\nu}_{t+1} = \omega\hat{\nu}_t + (1-\omega)\nu_t \tag{1.3}$$

where the measured disturbance, ν_t, is found by rearranging (1.1),

$$\nu_t = y_t - bu_t, \tag{1.4}$$

b is the estimate of the process gain (assumed for the moment to be equivalent to the true process gain, β), and ω is the estimate of θ. These parameters are determined

a priori through a designed experiment (DOE) or historical data. The result in (1.3) shows that the MMSE estimate for the disturbance is a weighted average of the current measured disturbance and the previous estimate of the disturbance. In practice, the EWMA tuning parameter, ω, is chosen by a control engineer to govern how quickly process data is discounted. When ω is chosen to be close to one, the estimate of ν is updated very slowly and when ω is zero only the most recent process measurement is considered in estimating ν.

After the EWMA filter has estimated the disturbance, a control law is used to determine the tool settings (or recipe) for the following run. In the unconstrained single input-single output case, the recipe is determined through simple model inversion,

$$u_{t+1} = \frac{T - \hat{\nu}_{t+1}}{b} \tag{1.5}$$

where T is the process target. When the MMSE disturbance estimate in (1.3) is paired with the control law in (1.5), and the estimated gain is equivalent to the true process gain, $b = \beta$, then the result is an MMSE EWMA controller. A process which is controlled by a MMSE EWMA controller has a closed loop variance equivalent to the variance of the noise, σ_0^2 [8].

The MIMO control law is somewhat more complicated as b may be non-square so that an inverse is not attainable. The MIMO control law, therefore, can take several different forms depending on the objective function of the optimization problem. A few of the common unconstrained MIMO control laws are listed below where the controller objective functions are followed by their closed form solutions.

1. Minimize the sum of the manipulated variables squared subject to the model hitting the process target [45]:

$$\min_{u_{t+1}} J = u_{t+1}^T u_{t+1} \quad s.t. \quad T = bu_{t+1} + \hat{\nu}_t$$

$$u_{t+1} = b^T(bb^T)^{-1}(T - \hat{\nu}_t) \tag{1.6}$$

2. Minimize the sum of the change in the manipulated variables squared subject to the model hitting the process target [63]:

$$\min_{u_{t+1}} J = \Delta u_{t+1}^T \Delta u_{t+1} \quad s.t. \quad T = bu_{t+1} + \hat{\nu}_t$$

$$u_{t+1} = b^T(bb^T)^{-1}(T - \hat{\nu}_t) + (I - b^T(bb^T)^{-1}b)u_t \tag{1.7}$$

where Δu_{t+1} is the change in controller input between successive runs, $\Delta u_{t+1} = u_{t+1} - u_t$.

3. Minimize the sum of squares deviation from the process target [17]:

$$\min_{u_{t+1}} J = (T - \hat{y}_{t+1})^T(T - \hat{y}_{t+1}) \quad s.t. \quad \hat{y}_{t+1} = bu_{t+1} + \hat{\nu}_t$$

$$u_{t+1} = (b^T b)^{-1} b^T (T - \hat{\nu}_t) \tag{1.8}$$

4. Model predictive control (MPC) formulation [12, 45]:

$$\min_{u_{t+1}} J = (T - \hat{y}_{t+1})^T Q (T - \hat{y}_{t+1}) + u_{t+1}^T R u_{t+1} + \Delta u_{t+1}^T S \Delta u_{t+1}$$

$$s.t. \qquad \hat{y}_{t+1} = bu_{t+1} + \hat{\nu}_t$$

$$u_{t+1} = (b^T Q b + R + S)^{-1}(S u_t + b^T Q (T - \hat{\nu}_{t+1})) \tag{1.9}$$

The first two control laws are applied when the number of inputs exceeds the number of outputs. In this case there are an infinite number of control inputs that will bring the process to the expected target, T. It follows that an objective function must be defined to establish a criteria with which to choose the 'best' controller input. The first objective function is to minimize the sum of squared controller input. The solution is similar to the single input-single output (SISO) case in (1.5) where the inverse of b is replaced with the right pseudo-inverse of b, $b^T(bb^T)^{-1}$.

The second objective function is the minimization of the sum of squared *change* in the controller input. The solution to this objective function is similar to the first solution in (1.6) with an additional term, $(I - b^T(bb^T)^{-1}b)u_t$, which penalizes the deviation from the previous controller input.

The third control law is used when the number of outputs exceeds the number of inputs. In this case there exists no controller inputs that will bring the process to the expected target. The objective function in this case is to minimize the sum of the squared deviation of the expected output from the target. The solution is similar to the solution in (1.6) where the left pseudo-inverse of b, $(b^Tb)^{-1}b^T$, is used in place of the right pseudo-inverse.

The final control law, shown in (1.9) is the most general of the four controller objective functions as the objective is to find the optimal balance between missing the process target, the absolute controller input, and the change in the controller input from the previous run. The control law in (1.9) can be made to return equivalent results as the other three control laws by selecting the appropriate values of Q, R, and S [12]. For example, by choosing a large Q, $R = I$ and $S = 0$, the solution

to the MPC formulation in (1.9) is identical to the solution in (1.6).

1.1.1.2 The Double EWMA Controller

It is often the case in semiconductor manufacturing that tools drift too rapidly from one run to the next for a process disturbances to be considered locally constant. For example, if a drift term, δ, is added to the disturbance expression in (1.2) according to

$$\nu_t = \nu_{t-1} - \theta\varepsilon_{t-1} + \delta + \varepsilon_t, \tag{1.10}$$

then the EWMA filter is not capable of adjusting to the process drift fast enough to prevent an offset in the control error. At steady state, the EWMA controller has a mean squared error of [15]

$$\lim_{t\to\infty} E\left[(T - y_t)^2\right] = \sigma_0^2 + \frac{(1-\omega)\frac{\beta}{b}\sigma_0^2}{2-(1-\omega)\frac{\beta}{b}} + \left(\frac{\delta}{(1-\omega)\frac{\beta}{b}}\right)^2. \tag{1.11}$$

To compensate for the steady state offset in (1.11), Butler and Stefani introduced the predictor corrector controller (PCC) [11]. The PCC incorporates an an estimate of the drift term, $\hat{\delta}$, in the disturbance model to compensate for the deterministic drift. This disturbance model has the form

$$\hat{\nu}_t = \hat{\nu}_{t-1} + \hat{\delta}_{t-1} \tag{1.12}$$

where we see that the disturbance estimate, $\hat{\nu}$, drifts by a factor of $\hat{\delta}$ for each process run. The function of the PCC filter is to recursively estimate both the intercept and the drift terms from process data. The updating equations for the estimates of the process intercept and drift are

$$\begin{aligned} \hat{\nu}_t &= \omega_1\hat{\nu}_{t-1} + (1-\omega_1)(y_t - bu_t) \\ \hat{\delta}_t &= \omega_2\hat{\delta}_{t-1} + (1-\omega_2)(y_t - bu_t - \hat{\nu}_{t-1}). \end{aligned} \tag{1.13}$$

It follows that by combining the one step ahead estimates for ν and δ in (1.13), the one step ahead prediction for the process disturbance becomes

$$\hat{\nu}_{t+1} = \hat{\nu}_t + \hat{\delta}_t. \tag{1.14}$$

Butler and Stefani demonstrate that the PCC is capable of compensating for a deterministic drift by applying the controller to a drifting polysilicon gate etch process. They show that the PCC converges to the true disturbance as the number of runs approaches infinity,

$$\lim_{t\to\infty}(\hat{\nu}_{t-1} + \hat{\delta}_t) = (\nu_{t-1} + \delta). \tag{1.15}$$

Nevertheless, Chen and Guo [19] show that the disturbance estimates from the PCC algorithm are biased and it is, therefore, difficult to assign a meaning to the asymptotically convergent values of the state estimates,

$$\lim_{t\to\infty}(\hat{\nu}_t) = \nu_0 + \delta t - \frac{\delta}{\omega_1}$$

$$\lim_{t\to\infty}(\hat{\delta}_t) = \frac{\delta}{\omega_1}$$

where ν_0 is the initial condition of the disturbance. The run-to-run controller introduced by Chen and Guo in [19], which they termed the double exponentially weighted moving average (double-EWMA) controller, has a similar format to Butler and Stefani's PCC algorithm. The parameter estimates of the double-EWMA observer, however, are unbiased and, therefore, have a more straightforward interpretation. The updating equations for Chen and Guo's double-EWMA algorithm are

$$\begin{aligned} \hat{\nu}_t &= \omega_1(\hat{\nu}_{t-1} + \hat{\delta}_{t-1}) + (1-\omega_1)(y_t - bu_t) \\ \hat{\delta}_t &= \omega_2\hat{\delta}_{t-1} + (1-\omega_2)(y_t - bu_t - \hat{\nu}_{t-1}) \end{aligned} \tag{1.16}$$

and the asymptotic estimates are

$$\begin{array}{c} \lim_{t\to\infty}(\hat{\nu}_t) = \nu_0 + \delta t \\ \lim_{t\to\infty}(\hat{\delta}_t) = \delta. \end{array} \tag{1.17}$$

A double-EWMA filter can also be derived in terms of discounted least squares regression. In this case, ordinary least squares regression is applied with discounted data to the linear drift model in (1.12). Abraham and Ledolter show that the updating equations for the recursive least squares (RLS) approach to the double-EWMA controller is [1]

$$\begin{array}{c} \hat{\nu}_t = \hat{\nu}_{t-1} + \hat{\delta}_{t-1} + (1-\omega^2)(y_t - \hat{y}_{t|t-1}) \\ \hat{\delta}_t = \hat{\delta}_{t-1} + (1-\omega)^2(y_t - \hat{y}_{t|t-1}) \end{array} \tag{1.18}$$

where $\hat{y}_{t|t-1}$ is the one step ahead forecast of y calculated at time, $t-1$,

$$\hat{y}_{t|t-1} = bu_t + \hat{\nu}_{t-1} + \hat{\delta}_{t-1}. \tag{1.19}$$

Wang and Qin showed that the single-tuning-parameter RLS derivation of the double-EWMA controller in (1.18) is equivalent to the two-tuning-parameter double-EWMA formulation in (1.16) when [66]

$$\begin{array}{c} \omega_1 = \omega^2 \\ \omega_2 = 2\omega - \omega^2. \end{array}$$

Although the double-EWMA formulations in (1.16) and (1.18) are similar in their formulations and offer unbiased estimates of the disturbance parameters, there is a fundamental difference in their derivations. The double-EWMA formulation in (1.18) is derived by fitting a linear model to discounted data. In this case, the process engineer is responsible for simply choosing how quickly data should

be discounted. The double-EWMA formulation in (1.16), on the other hand, requires the process engineer to choose how quickly the model parameters themselves should be discounted. The elimination of the second tuning parameter allows for a simpler (and somewhat less arbitrary) derivation of the double-EWMA run-to-run controller. For this reason, the double-EWMA formulation in (1.18) will be used exclusively in this dissertation when referring to the double-EWMA controller.

Both the SISO and MIMO control laws for the double-EWMA controller are analogous to the EWMA control laws shown in (1.6) through (1.9). The only difference is the disturbance estimate: where the EWMA controller assumes a constant disturbance, $\hat{\nu}_t$, the double-EWMA controller assumes a drifting disturbance, $\hat{\nu}_t + \hat{\delta}_t$. For example, the control law in (1.6) would be

$$u_{t+1} = b^T(bb^T)^{-1}(T - \hat{\nu}_t - \hat{\delta}_t) \tag{1.20}$$

for the double-EWMA controller.

1.1.1.3 Other Run-to-Run Controllers

Several additional run-to-run control algorithms have been proposed in the literature but are less commonly utilized in industry. The optimizing adaptive quality controller (OAQC) introduced by Del Castillo and Yeh [18] utilizes Åström and Wittenmark's recursive least squares algorithm to estimate the parameters of an assumed process model [4]. Although both the EWMA and double-EWMA controllers are recursive least squares algorithms, the OAQC algorithm is unique in that, in addition to estimating the disturbance(s), the process gain, b, is recursively estimated. Although the OAQC algorithm has the obvious advantage of being able

to adapt to a wide range of operating regions, the controller suffers from the requirement of needing constant excitation to accurately estimate the gain matrix. If the covariance of the data becomes rank deficient then the closed loop process will become unstable. In addition, since the algorithm utilizes simple model inversion as the control law, if the gain estimate approaches zero then the control input will approach infinity.

The age-based double-EWMA controller is similar to the double-EWMA controller except the process is assumed to have a deterministic drift by *time* instead of by process runs [19]. This controller is unique in that it compensates for equipment utilization. Chen and Guo argue that the periods of inactivity on a process tool are commonly caused by the tool running different products or different manufacturing steps. That is, the tool is not idle between updates to the disturbance state, but rather the tool continues to operate and, therefore, drift from the process target(s). Chen and Guo confirm this hypothesis by demonstrating larger drifts between runs when there is a larger downtime between process runs.

Several other run-to-run controllers have been proposed in the literature as well. Zafiriou, Adomaitis, and Gattu's gradient optimization method searches for directions of process improvement with successive process runs [69]. Baras and Patel's robust run-to-run controller uses a dynamic program to solve an H_∞ optimal trajectory which minimizes the worst case process performance for a process with metrology delay and model mismatch [5]. Smith and Boning's neural network based controller uses a neural network to map process recipes to process outputs [60]. This neural input-output relationship is used to determine the optimal recipes to reach

a desired process target.

Comparisons of different run-to-run control algorithms can be found in [13, 48] and an extensive literature review of existing run-to-run control algorithms can be found in [6, 16].

1.1.2 Stability of Run-to-Run Controllers

The first work concerning the conditions which determine the stability of a process under closed-loop feedback control can be attributed to Maxwell's analysis of the Watt governor – a mechanical device used to control the speed of a steam engine [23, 46]. Using linearized differential equations, Maxwell showed that the governor was stable when all of the roots of the characteristic equation are in the open LHP. In addition, he derived the stability conditions for up to third order polynomials. Routh and Hurwitz later independently developed numerical techniques for determining if all of the roots of a polynomial are in the LHP without having to actually solve for the roots of the polynomial [36, 53].

After these pioneering works on closed-loop stability, the development of advanced control techniques led to the development of more advanced stability analysis techniques. One topic of particular interest for this dissertation concerns the stability of systems with uncertainty in the process model. In a seminal work on the stability of uncertain processes, Palmor writes,

> The sensitivity of the stability and the performance of a controlled system to parameter variations and plant ignorance is probably the most important yardstick by which any practical controller is measured. This

> is especially true in process control where complex nonlinearities and time varying processes are represented in the design by linear constant low order models [50].

Clearly, because of the use of inaccurate low-order models and the nonlinear nature of semiconductor manufacturing tools, run-to-run controllers are subject to closed-loop instability.

The first study on the conditions for stability of the SISO EWMA controller was published by Ingolfsson and Sachs shortly after their first work introducing run-to-run control for semiconductor manufacturing [37]. They noted that a process will become unstable when the input-output relationship between the tool recipes and quality measurements are not accurately estimated. The allowable range of model mismatch that a process can have and still maintain asymptotic stability was shown to be

$$0 < \xi(1 - \omega) < 2 \tag{1.21}$$

where ξ is defined as the ratio between the true process gain, β, and the estimated process gain, b,

$$\xi = \frac{\beta}{b}. \tag{1.22}$$

The result in (1.21) illustrates that the EWMA controller is robust for a wide range of model mismatch. As long as the estimated process gain is greater than half of the true process gain, then the closed loop system will be asymptotically stable.

The stability of the MIMO EWMA controller was first examined by Kosut *et al* [43]. Although their stability boundary was later shown to be incorrect, they

correctly asserted that the stability of the MIMO EWMA controller is governed by the eigenvalues of the model-mismatch matrix, $\boldsymbol{\xi} = \beta b^T(bb^T)^{-1}$. Tseng *et al* later extended the work of Kosut *et al* to show the stable region the formulation of the MIMO EWMA controller version in (1.7) [63]. The stability region for this particular MIMO EWMA controller is

$$\frac{\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a}}{\mathrm{a}^2 + \mathrm{b}^2} < \omega < 1 \qquad \forall \lambda_j = \mathrm{a} + \mathrm{b}i \tag{1.23}$$

where λ_j are the eigenvalues of the model mismatch matrix.

The effect of metrology delay on the stability of run-to-run controllers was first studied by Adivikolanu and Zafiriou [2, 3]. They utilized an internal model control approach to derive the stability region of the SISO EWMA controller with a delay of one run,

$$0 < \xi < \frac{2 - \omega}{1 - \omega}. \tag{1.24}$$

A numerical method was then introduced for determining the stability region of the EWMA controller for processes with longer metrology delays.

There has been similar interest in studying the closed-loop stability conditions for the double-EWMA controller family. The stability of the SISO PCC was first studied by Del Castillo in [14] and this work was later extended to the MIMO PCC by Del Castillo and Rajagopal [17]. The SISO PCC stability region was shown to be

$$|1 - 0.5\xi(\omega_1 + \omega_2) \pm 0.5z| < 1 \tag{1.25}$$

where

$$z = \sqrt{\xi^2(\omega_1 + \omega_2)^2 - 4\omega_1\omega_2\xi}.$$

This result states that the poles of the closed loop system must be inside the unit circle (a true statement for any process) and offers little insight as to the true closed loop stability conditions. Castillo and Rajagopal reported that the MIMO PCC stability region is

$$|1 - 0.5\xi_{jj}(\omega_1 + \omega_2) \pm 0.5z| < 1 \qquad \forall j \tag{1.26}$$

where

$$z = \sqrt{\xi_{jj}^2(\omega_1 + \omega_2)^2 - 4\omega_1\omega_2\xi_{jj})}$$

and ξ_{jj} are the diagonal elements of $\boldsymbol{\xi}$. It will be shown in Chapter 3 through a counterexample that this solution is incorrect and the correct solution to the double-EWMA stability region is derived there.

1.1.3 Performance of Run-to-Run Controllers

Previous work on the closed-loop performance of run-to-run controllers has focused on two key areas: transient behavior and long term behavior. The transient behavior is concerned with how quickly the controller responds to a process disturbance and is typically investigated through simulations [15]. The transient response can also be characterized in terms of the closed-loop eigenvalues of the process. Kosut *et al* defined closed-loop performance in terms of an equivalent time constant [43]. This, however, was a somewhat limited analysis as only the SISO case was considered with an EWMA tuning parameter of $\omega = 0$. Adivikolanu and Zafiriou showed in [2, 3] that techniques similar to those used to study closed-loop stability can be used to derive closed-loop performance criteria. However, they too only derived the performance boundaries for the SISO case and did not consider the

effects of metrology delay.

Beyond showing that the expected output converges on the process target, it is difficult to characterize the long-term performance of a run-to-run controller. As mentioned previously, Box and Jenkins showed that an EWMA controller is a MMSE controller if a disturbance follows an IMA{1,1} time series [8]. It follows that if the disturbance does not follow an IMA{1,1} time series, then the EWMA controller is not an MMSE controller and the long-term variance will be greater than the variance of the disturbance noise, σ_0^2. Del Castillo derived the long-term variance of the EWMA controller for the following scenarios [15].

- An IMA{1,1} disturbance with $\omega \neq \theta$
- An IMA{1,1} disturbance with $\omega \neq \theta$ and a deterministic drift, δ
- A white noise disturbance
- A white noise disturbance with a deterministic trend

For the above scenarios, the closed loop variance is

$$\sigma^2 = \frac{1 + \theta^2 - 2(1 - \xi(1-\omega))\theta}{\xi(1-\omega)(2 - \xi(1-\omega))}\sigma_0^2 + \left(\frac{\delta}{(1-\omega)\xi}\right)^2. \tag{1.27}$$

Nevertheless, like studies on the transient performance, the majority of the published results on the long term performance of run-to-run controllers are based on simulations (for example [33, 41, 48, 63, 64, 70]).

1.2 Research Objectives

The overall objective of this research is to further investigate the closed loop stability and performance characteristics of the EWMA and double-EWMA run-to-run controllers. Although much research has been performed on this topic to date, it is clear that the work of Ingolfsson and Sachs, Del Castillo and Rajagopal, Tseng *et al*, Adivikolanu and Zafiriou, and others lack a comprehensive study of the effects of metrology delay in the stability and performance of run-to-run controllers. Therefore, this research will address the following topics:

1. Derive the closed loop stability criteria for both SISO and MIMO EWMA and double-EWMA run-to-run controlled systems with metrology delay and parametric uncertainty in the process gain.

2. Introduce an accurate approximation of the closed loop stability boundaries to reduce the complexity of the analytical solution for systems with extended metrology delays.

3. Extend the stability analysis to consider the closed loop *performance* criteria of both SISO and MIMO run-to-run controllers with and without metrology delays.

4. Derive the optimal double-EWMA tuning parameters in terms of maximizing the possible model mismatch while meeting a predetermined closed-loop performance criteria.

Table 1.1: The stability of run-to-run controllers: the state-of-the-art and topics covered in this dissertation

	EWMA		Double-EWMA	
Delay	SISO	MIMO	SISO	MIMO
0	[37]	[63]	[3, 14]	[15] [a]
1	[3] [b]	√	√	√
2	√	√	√	√
⋮	⋮	⋮	⋮	⋮
10 [c]	√	√	√	√

[a] Erroneous derivation, the correct solution is derived in this dissertation

[b] A numerical technique is provided for delays of greater than one run

[c] Empirical approximations are introduced for delays greater than two runs

Table 1.2: The performance of run-to-run controllers: the state-of-the-art and topics covered in this dissertation

	EWMA		Double-EWMA	
Delay	SISO	MIMO	SISO	MIMO
0	[3]	[63]	[14]	√
1	√	√	√	√
2	√	√	√	√

The state-of-the-art in the analysis of run-to-run controller stability and performance is summarized in Tables 1.1 and 1.2, respectively. Also shown in Tables 1.1 and 1.2 are the works that are presented in this dissertation where the open topics that are addressed in this work are marked with a "√". It is clear that there are many open topics within the field of run-to-run control stability and performance and it is the goal of this work to addresses these topics.

1.3 Dissertation Summary

The outline of the dissertation is as follows. The closed loop stability criteria of SISO and MIMO EWMA controllers with metrology delay and plant-model mismatch (i.e., when $b \neq \beta$) is derived in Chapter 2. Here, analytical expression for the closed loop stability of delayed processes are derived and a sufficient condition and a necessary condition for stability of processes with longer metrology delays is provided. An analogous stability analysis of the double-EWMA controller is then provided in Chapter 3.

As metrology delay increases, the analytical solutions to the stability bounds become increasingly complex. Therefore, a numerical approximation technique for the stability regions of the EWMA and double-EWMA controllers are provided in Chapter 4.

The performance of the EWMA and double-EWMA run-to-run controllers with metrology delay are discussed in Chapters 5 and 6, respectively. Here, instead of simply constraining the closed-loop poles to be inside the unit circle, a minimum closed-loop performance is guaranteed by constraining the eigenvalues to a minimum performance region (a circle with a radius smaller than unity). In addition, the double-EWMA tuning parameter which allows for the largest model mismatch while meeting a predetermined closed-loop performance is derived in Chapter 6.

Finally some conclusions and recommendations for future work in the field of run-to-run controller stability and performance are provided in Chapter 7.

Chapter 2

The Stability of the EWMA Controller

In semiconductor manufacturing, it is rarely possible to perform quality measurements on a product before subsequent process operations are performed because of throughput requirements and constraints on metrology capacity. This delay between product manufacturing and product metrology coupled with inaccurate process models can lead to process instabilities. This chapter investigates the robust stability requirements of semiconductor manufacturing processes that are controlled with an EWMA controller. Both SISO and MIMO processes are considered as well as the effect of metrology delay on the robust stability conditions. It is shown in this chapter that a sufficient requirement for stability of the SISO process is that the true process gain must be between zero and two times the process model gain, regardless of the length of metrology delay. Likewise, a sufficient condition for stability of the MIMO case is that all of the eigenvalues of the model mismatch matrix fall inside a circle centered at {1,0} on the complex plane and with unit radius.

2.1 Stability of the SISO EWMA Controller

In this section the stability conditions for the SISO EWMA controller will be derived both with and without metrology delay. The stability condition for a

process without metrology delay is considered in Section 2.1.1 and the conditions for stability of a system with metrology delay is considered in Section 2.1.2. An illustrative simulation is provided in Section 2.1.3 to demonstrate the importance of considering metrology delay in tuning an EWMA controller.

2.1.1 SISO Stability Without Metrology Delay

The closed loop response of a system being controlled by a SISO EWMA controller can be derived by combining the process model in (1.1) and the disturbance estimates in (1.3) with the control law in (1.5),

$$\hat{\nu}_{t+1} = \left(1 - \frac{\beta}{b}(1-\omega)\right)\hat{\nu}_t + (1-\omega)\left(\frac{\beta}{b}T + \nu + \varepsilon_{t+1}\right). \tag{2.1}$$

Without loss of generality we can set $T = 0$ for a constant target setpoint so that (2.1) can be rewritten

$$\hat{\nu}_{t+1} = (1 - \xi(1-\omega))\,\hat{\nu}_t + (1-\omega)(\nu + \varepsilon_{t+1}) \tag{2.2}$$

where the model mismatch has been defined as $\xi = \beta/b$. It is easy to see from (2.2) that the estimate of the process intercept, $\hat{\nu}$, will be stable when $-1 < (1 - \xi(1-\omega)) < 1$. Solving this inequality yields the stability region for the SISO EWMA controller,

$$0 < \xi < \frac{2}{1-\omega}, \tag{2.3}$$

which is equivalent to the result obtained in [37]. This stability region in shown in Figure 2.1.

Figure 2.1 shows that as the model mismatch increases, a larger ω is necessary for the process to remain stable. We also see in Figure 2.1 that a necessary

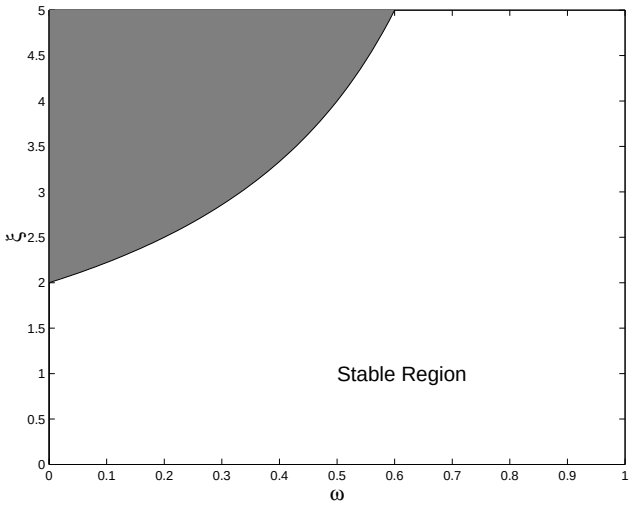

Figure 2.1: Stability region for the SISO EWMA controller without metrology delay

condition for stability is that the estimated process gain must have the same sign as the true process gain, (i.e, $\xi > 0$). It follows that a sufficient condition for stability is that the estimated process gain is larger than the true process gain for any $0 \leq \omega < 1$. That is, by over estimating the gain, closed-loop stability is guaranteed.

At steady state (i.e., when $\hat{\nu}_{t+1} = \hat{\nu}_t$) (2.2) becomes

$$\xi \hat{\nu}_{t+1} = \nu + \varepsilon_{t+1} \tag{2.4}$$

which indicates that $\hat{\nu}_t$ is a biased estimate of ν when $\xi \neq 1$. Therefore, the presence of model mismatch will result in an inaccurate estimate of the process disturbance. However, combining (1.3) and (1.5) leads to

$$u_{t+1} = u_t - \frac{1-\omega}{b} y_t \tag{2.5}$$

which is an integral controller and thus the process will have no offset at steady

state. Therefore, as long as the system is stable, the output will converge to the process target.

2.1.2 SISO Stability With Metrology Delay

As mentioned previously, because of demands on semiconductor throughput and constraints in metrology capacity, measurements can rarely be performed on a batch of wafers before the next batch of wafers is processed. To account for this metrology lag, the d step ahead forecast must be used in estimating the process disturbance. By assuming that the disturbance follows an IMA{1,1} time series, the $d+1$ step ahead forecast is

$$\hat{\nu}_{d+1|t} = \hat{\nu}_{d|t} = \ldots = \hat{\nu}_{t+1|t} \tag{2.6}$$

where the notation $\hat{\nu}_{d+1|t}$ refers to the $d+1$ step ahead forecast estimated at time, t. Using the $d+1$ forecast, the updating equations for the intercept estimate in (1.3) is rewritten,

$$\hat{\nu}_{t+1} = \omega\hat{\nu}_t + (1-\omega)(y_{t-d} - bu_{t-d}) \tag{2.7}$$

where d is the number of runs between product manufacturing and product metrology. By combining the process model (1.1) and the control law (1.5) with the EWMA filter (2.7), the closed loop equation is written as

$$\hat{\nu}_{t+d+1} = \omega\hat{\nu}_{t+d} + (1-\omega)(1-\xi)\hat{\nu}_t + (1-\omega)\nu + \varepsilon_{t+1}. \tag{2.8}$$

From (2.8) we see that the updating equation for $\hat{\nu}$ is a function of the previous calculation of $\hat{\nu}$ and $\hat{\nu}$ from the $(t-d)^{th}$ run. Taking the z-transform of (2.8) yields

the characteristic equation,

$$f(z) = z^{d+1} - z^d\omega - (1-\xi)(1-\omega) = 0, \tag{2.9}$$

so that the closed loop system is stable when all of the roots of (2.9) are inside the unit circle. To find the conditions that meet this stability requirement, the inside of the unit circle is mapped to the left half plane (LHP) using the bilinear transformation,

$$z = \frac{1+w}{1-w}. \tag{2.10}$$

The bilinear transform allows us to reduced the problem to finding the conditions where the roots for w are in the LHP (a much simpler problem). It is worth noting here that techniques exist to directly solve the discrete stability problem [22, 30, 40, 56]. However the complexity of the solution is greatly reduced through this mapping step. This simplification will become more apparent as processes with longer metrology delay are considered.

The procedure for determining EWMA closed-loop stability conditions will be demonstrated by looking at cases when $d = 1$ and $d = 2$.

2.1.2.1 The Stability of a Process With a Delay of One Run

After the bilinear transformation step in (2.10), the characteristic equation of a process with a delay of one run becomes

$$f(w) = (2\omega + \xi - \omega\xi)w^2 + (4 - 2\omega - 2\xi + 2\omega\xi)w + \xi(1-\omega) = 0. \tag{2.11}$$

The roots of (2.11) are in the LHP when all of the coefficients on w have the same sign so that the system is stable when the following three conditions are met,

1. $2\omega + \xi - \omega\xi > 0$

2. $4 - 2\omega - 2\xi + 2\omega\xi > 0$

3. $\xi(1 - \omega) > 0$.

Solving this system of inequalities yields three potential candidates for the active stability boundaries:

1. $\xi > -\frac{2\omega}{1-\omega}$

2. $\xi < \frac{2-\omega}{1-\omega}$

3. $\xi > 0$.

The first boundary is less than the third boundary for all values of $0 \leq \omega \leq 1$ so that the first boundary is not an 'active' boundary. This leaves the second and third boundaries as the active stability boundaries for the system: the second boundary acts as the upper boundary on the model mismatch and the third boundary acts as the lower boundary. It follows that the stable region of the EWMA controller with a delay of one run is

$$0 < \xi < \frac{2 - \omega}{1 - \omega}. \tag{2.12}$$

This solution is shown in Figure 2.2 and is equivalent to the solution in [2, 3].

2.1.2.2 The Stability of a Process With a Delay of Two Runs

By extending the delay to two runs the characteristic equation in (2.9) becomes

$$f(z) = z^3 - z^2\omega - (1 - \xi)(1 - \omega) = 0 \tag{2.13}$$

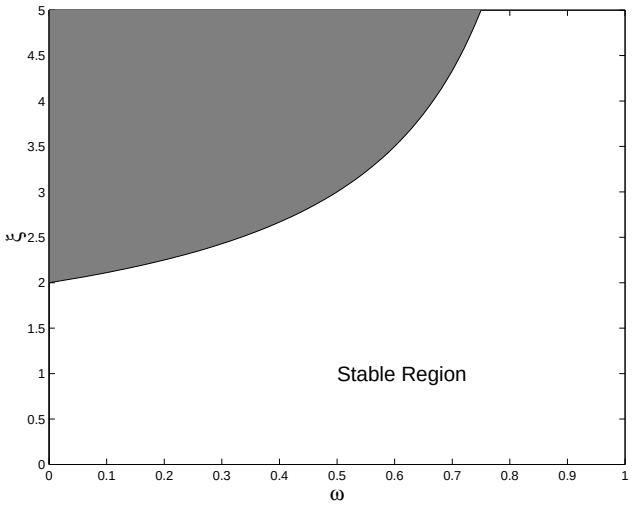

Figure 2.2: Stability region for the SISO EWMA controller with a delay of one run

and after mapping the inside of the unit circle to the LHP, the characteristic equation becomes

$$f(w) = (-\omega\xi + \xi - 2)w^3 + (3\omega z - 4\omega - 3\xi)w^2$$
$$-(3\omega\xi - 4\omega - 3\xi + 6)w - \xi(1-\omega) = 0. \tag{2.14}$$

The conditions where all of the roots are in the LHP are more complicated with systems with a metrology lag of more than one run. There are several techniques that can be used to find when a system has all of its roots in the LHP including Jury's test [40], the Routh array technique [53], the Routh-Hurwitz stability criterion [36], and the Lienard-Chipart stability criterion [44]. Although all four techniques will presumably arrive at the same stability region, the Routh-Hurwitz stability criterion is easily extendible to higher order systems and will be

used exclusively in this chapter and throughout the remainder of this dissertation.

The Routh-Hurwitz stability criterion states that a polynomial of the form,

$$q(s) = a_p s^p + a_{p-1} s^{p-1} + \cdots + a_1 s + a_0 = 0, \tag{2.15}$$

and with all real coefficients, has all of its roots in the LHP when all of the principal minors of the matrix,

$$H \equiv \begin{bmatrix} a_{p-1} & a_{p-3} & a_{p-5} & a_{p-7} & \cdots & \\ a_p & a_{p-2} & a_{p-4} & a_{p-6} & \cdots & \\ & a_{p-1} & a_{p-3} & a_{p-5} & \cdots & \\ & a_p & a_{p-2} & a_{p-4} & \cdots & \\ & & \vdots & \vdots & \ddots & \\ & & & & a_{p-2} & a_0 \end{bmatrix}, \tag{2.16}$$

are greater than zero. Further details on the Routh-Hurwitz stability criterion can be found in [36, 53, 57].

Applying the Routh-Hurwitz stability criterion on (2.14) yields

$$H = \begin{bmatrix} 3\omega\xi - 4\omega - 3\xi & -\xi(1-\omega) & 0 \\ -\omega\xi + \xi - 2 & -3\omega\xi + 4\omega + 3\xi - 6 & 0 \\ 0 & 3\omega\xi - 4\omega - 3\xi & -\xi(1-\omega) \end{bmatrix} \tag{2.17}$$

and solving the system of inequalities which guarantees all of the principal minors of H are greater than zero results in the following candidates for the stability boundaries:

1. $\xi > \frac{2-3\omega-\sqrt{\omega^2+4}}{2(1-\omega)}$
2. $\xi < \frac{2-3\omega+\sqrt{\omega^2+4}}{2(1-\omega)}$
3. $\xi < \frac{2}{1-\omega}$

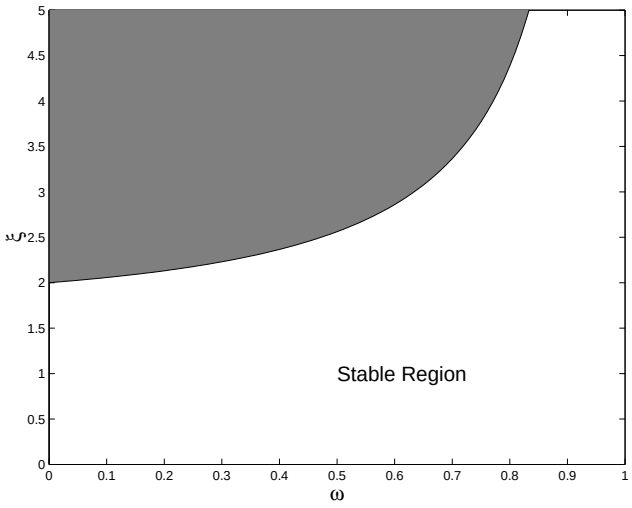

Figure 2.3: Stability region for the SISO EWMA controller with a delay of two runs

4. $\xi > 0$.

Like the case with a delay of one run, the first stability boundary is negative for $\omega > 0$. Since this result is less than the fourth boundary, the first boundary is not an active boundary. The second boundary is less than the third boundary in the range $\omega > 0$ so that the second boundary is eliminated as an active boundary. This leaves boundaries two and four as the active boundaries. It follows that the stability region for the EWMA controller with a delay of two runs is,

$$0 < \xi < \frac{2 - 3\omega + \sqrt{\omega^2 + 4}}{2(1 - \omega)}. \tag{2.18}$$

This is shown in Figure 2.3.

By comparing Figures 2.1 through 2.3, we can see that increasing metrology delay decreases the size of the stability region. That is, for a given model mis-

match, ξ, it is necessary to use a less aggressive controller to guarantee stability when the metrology lag is increased from one run to two. We can also see a pattern of increasing complexity in the stability boundary expressions as the length of the metrology delay increases. To keep from having to evaluate the analytical expressions for increasing metrology delays, a sufficient condition for stability is derived in the following section.

2.1.2.3 The Stability of Systems With Large Metrology Delays

The procedure used to find the stable region of the system with a metrology lag of two runs can be extended to find the stable regions for processes with longer delays. Nevertheless, by looking at the two extremes of the EWMA tuning parameter, two general remarks can be made concerning EWMA controlled processes with any length metrology lag. First, we see that the stability boundary approaches infinity as the EWMA tuning parameter approaches unity. This is equivalent to turning the controller off (as the parameter estimates are never updated) and the system is stable as long as the disturbances are bounded. Next, we see that a system becomes less stable as the forgetting factor, ω, gets closer to zero. This could equivalently be stated that the minimum upper bound is reached when $\omega = 0$. Therefore, a sufficient condition for stability can be derived by considering a system with a forgetting factor, $\omega = 0$ (i.e., the worst case closed-loop scenario). Setting $\omega = 0$, the characteristic equation in (2.9) becomes

$$f(z) = z^{d+1} - (1 - \xi) = 0 \tag{2.19}$$

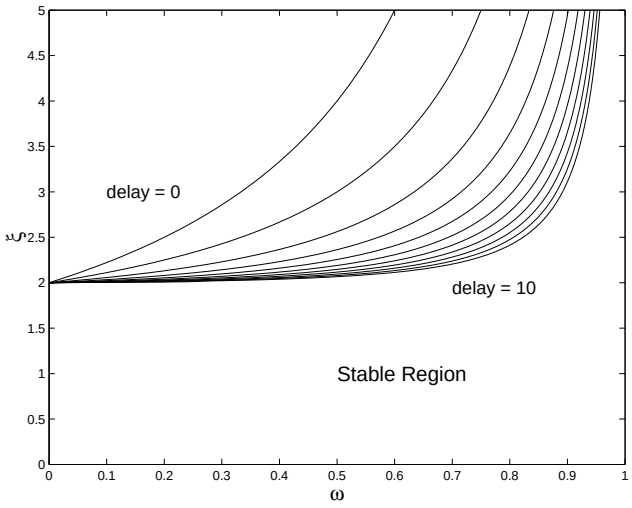

Figure 2.4: Stability region for the SISO EWMA controller with large metrology delays

so that the roots are all inside the unit circle when

$$0 < \xi < 2. \tag{2.20}$$

This leads to the following theorem:

Theorem 2.1.1. *A sufficient condition for stability of the SISO EWMA controller for a delay of any length is that the model mismatch, ξ, is between zero and two.*

It follows that by overestimating the process gain, the asymptotic stability of the closed loop system can be guaranteed. The stable regions for delays of up to ten runs are shown in Figure 2.4.

Figure 2.4 shows that, irrespective of the length of the metrology delay, an EWMA controlled process will be stable when the following criteria are met: the estimated gain, b, must have the same sign as the true process gain and the estimated

gain must be greater than half of the true process gain. Figure 2.4 also shows that the range of allowable gain estimates increases as the EWMA tuning parameter gets larger, (i.e., as the controller becomes less aggressive). As the tuning parameter approaches unity we see that the only constraint for stability is that the estimated gain must have the same sign as the true gain. This, however, says nothing of the closed loop performance, as a tuning parameter of unity is equivalent to an open loop process.

We also see in Figure 2.4 that this sufficient condition for stability is conservative for small metrology delays and then becomes less conservative with increasing metrology delay. For example, when $d = 2$ and $\omega = 0.5$, the necessary and sufficient condition for stability is $0 < \xi < 2.56$ which is 28% larger than the sufficient condition. Then, if the metrology delay is increased to $d = 10$, the necessary and sufficient condition for stability becomes $0 < \xi < 2.07$; only 3.3% larger than the sufficient condition.

2.1.3 Simulation of a SISO Process

In this section the closed loop stability of an EWMA controlled process is demonstrated using a simple simulation. Consider a system with $T = 0, \nu = 0, \beta = 7.0$ and with normally distributed noise with unit variance. Through a designed experiment or from process data, an approximation for the model gain is determined to be $b = 2.0$ so that the model mismatch is $\xi = 3.5$. According to the boundary regions shown in Figures 2.1 through 2.3, $\omega = 0.6$ will result in a stable system when there is no delay, a marginally stable system with a delay of one and an unstable

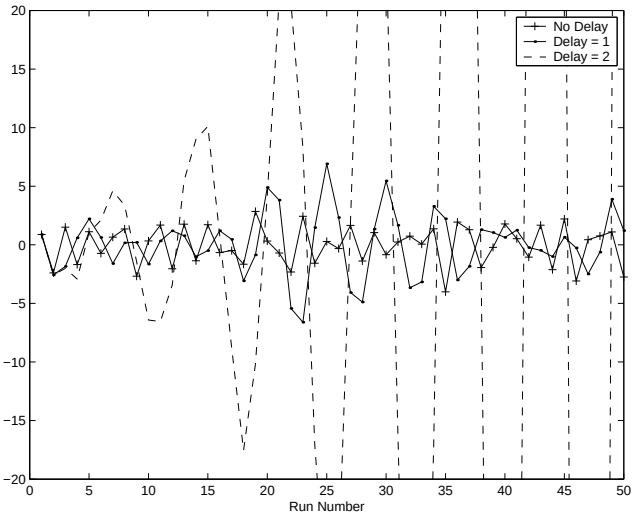

Figure 2.5: Simulation of a SISO EWMA controller with $\xi = 3.5$, $\omega = 0.6$, and delays of 0, 1, and 2 runs

system with a delay of two. Figure 2.5 shows 50 runs of this simulation. Here we see the importance of considering metrology delay implicitly in choosing the forgetting factor for an EWMA controller. A process that is stable at a given metrology lag can become unstable if the metrology delay is increased.

2.2 Stability of the MIMO EWMA Controller

In this section the stability conditions for the MIMO EWMA controller will be derived. The stability condition for a process without metrology delay is considered in Section 2.2.1 and that with metrology delay is considered in Section 2.2.2. An illustrative simulation is provided in Section 2.2.3.

2.2.1 MIMO Stability Without Metrology Delay

The closed loop realization of a process controlled with a MIMO EWMA controller is analogous to the SISO case,

$$\hat{\nu}_{t+1} = (I - \boldsymbol{\xi}(I - \omega))\,\hat{\nu}_t + (I - \omega)\,(\boldsymbol{\xi}T + \nu + \varepsilon_t)\,, \tag{2.21}$$

where the definition of the model mismatch matrix, $\boldsymbol{\xi}$, is dependent on the controller objective function:

1. Minimize the sum of squared control inputs

$$\boldsymbol{\xi} = \beta b^T (b b^T)^{-1} \tag{2.22}$$

2. Minimize the sum of squared change in control input

$$\boldsymbol{\xi} = \beta b^T (b b^T)^{-1} \tag{2.23}$$

3. Minimize the sum of squared deviation from target

$$\boldsymbol{\xi} = \beta (b^T b)^{-1} b^T \tag{2.24}$$

4. MPC formulation with $S = 0$

$$\boldsymbol{\xi} = (\beta - b)(b^T Q b + R)^{-1} b^T Q + I. \tag{2.25}$$

The derivations for these model mismatch matrices are shown in Appendix A.1. In addition, the tuning parameter, $\boldsymbol{\omega}$ in (2.22) through (2.25) is a diagonal matrix containing the EWMA tuning parameters for each of the n disturbance channels,

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{bmatrix}. \tag{2.26}$$

When deriving the EWMA filter from discounted recursive least squares (RLS) regression, process measurements are discounted at a rate of $(1-\omega)^t$ for each of the n disturbance channels [1]. In other words, the disturbance estimate is extracted from discounted *data* as opposed to filtering the updates of the n disturbance channels independently. It follows that, with discounted RLS regression, all of the disturbance channels are updated at the same rate so that $\omega_1 = \omega_2 = \ldots = \omega_n$ and only one tuning parameter is necessary for the MIMO EWMA controller,

$$\boldsymbol{\omega} = \begin{bmatrix} \omega & & \\ & \ddots & \\ & & \omega \end{bmatrix} = \omega I_n. \tag{2.27}$$

Using the discounted RLS interpretation of the MIMO EWMA controller, the MIMO system without metrology delay is stable when all of the eigenvalues of the transition matrix,

$$A = I - \boldsymbol{\xi}(I - \omega), \tag{2.28}$$

are inside the unit circle. Finding the stability conditions for the MIMO case is much more complicated than the SISO case since the transition matrix now has the same dimensions as the model mismatch matrix, $\boldsymbol{\xi}$. However, the complexity of the MIMO system can be reduced dramatically according to the following Lemma.

Lemma 2.2.1. *The model mismatch matrix can be replaced with its eigenvalues in determining the stability boundary of the MIMO system for* $\boldsymbol{\omega} = \omega I$*, that is,* $eig(A) = 1 - \lambda_j(1 - \omega)$ *where* λ_j *are the eigenvalues of* $\boldsymbol{\xi}$*.*

Proof: See Appendix A.2.

Although a complex model mismatch parameter is not physically possible for the SISO case, the eigenvalues of $\boldsymbol{\xi}$ for the MIMO case can be complex. Therefore, we have the following theorem for the stability boundary of the system without metrology delay.

Theorem 2.2.2. *Denoting $\lambda_j = \mathrm{a} + \mathrm{b}i$, the MIMO EWMA with $\boldsymbol{\omega} = \omega I$ controller is stable if and only if*

$$\frac{\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a}}{\mathrm{a}^2 + \mathrm{b}^2} < \omega < 1, \tag{2.29}$$

which is equivalent to the result in [63].

Proof: See Appendix A.3.

The complex stability regions for the EWMA controller without metrology delay are shown in Figure 2.6. Figure 2.6 shows that the stability of the EWMA controlled process shrinks as the complex part of λ increases. Unlike the SISO case, when the eigenvalues of $\boldsymbol{\xi}$ have a complex part, an unstable region develops as the real part, a, approaches zero. This indicates that, unlike the SISO case, overestimating the process gain can result in closed loop instability.

2.2.2 MIMO Stability With Metrology Delay

When metrology delay is considered, the closed-loop equation for the MIMO EWMA controller becomes

$$\hat{\nu}_{t+d+1} = \omega\hat{\nu}_{t+d} + (I - \omega)(I - \boldsymbol{\xi})\hat{\nu}_t + (I - \omega)\nu + \varepsilon_{t+1}. \tag{2.30}$$

Denoting $\lambda_1, \lambda_2, \ldots, \lambda_m$ as the eigenvalues of $\boldsymbol{\xi}$ and using the same argument as in the proof of Lemma 2.2.1, we see that the MIMO characteristic equation is analogous

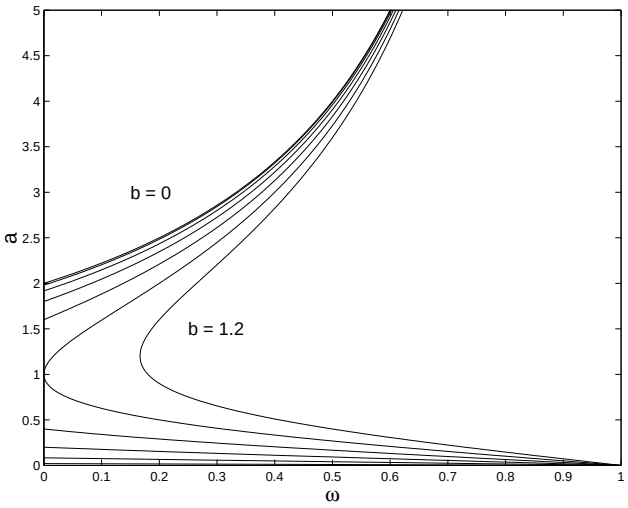

Figure 2.6: The complex stability boundaries for a system without metrology delay, b = 0 . . . 1.2 in increments of 0.2

the the SISO characteristic equation in (2.9) with respect to the eigenvalues of the model mismatch matrix,

$$f(z) = z^{d+1} - \omega z^d - (1-\omega)(1-\lambda_j) = 0, \tag{2.31}$$

for $j = 1, 2, \ldots m$.

Solving (2.31) for λ_j yields

$$\lambda_j = 1 - \frac{z^{d+1} - \omega z^d}{1-\omega} \tag{2.32}$$

where the z can be any value along the unit circle. By expressing z in polar coordinates, $z = e^{i\theta}$ for $-\pi \leq \theta \leq \pi$, the complex stability boundary can be plotted. This is shown in Figure 2.7 for $\omega = 0.5$ and delays of zero through three runs.

Figure 2.7 shows that the complex stability boundary in (2.32) is a spiral

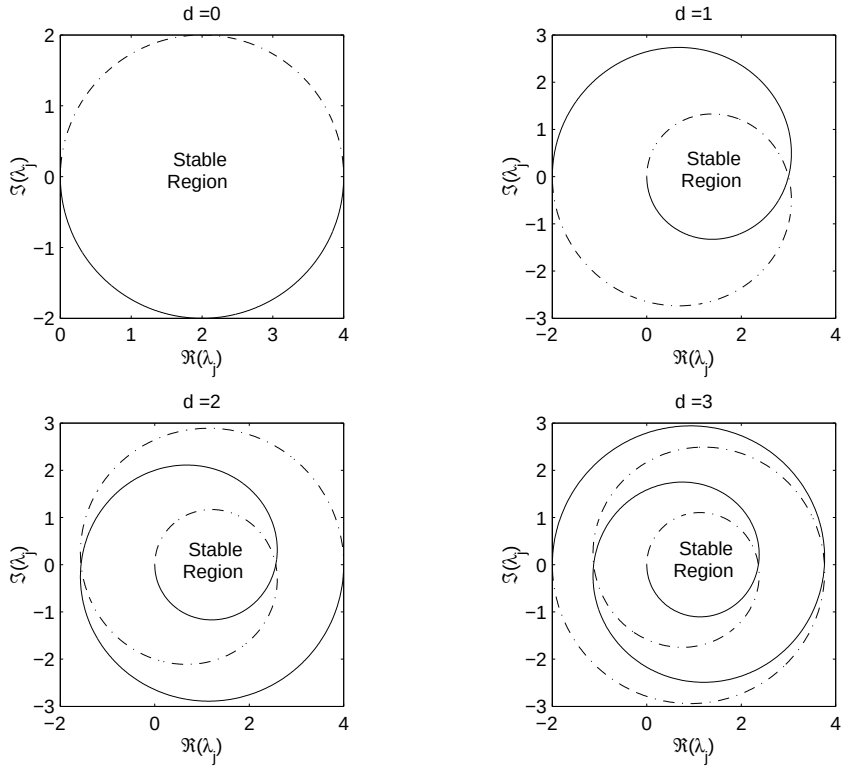

Figure 2.7: The complex stability region for $\omega = 0.5$. Starting from $\lambda_j = 0$, the complex stability bounds spiral out as θ approaches $\pm\pi$. The solid and dashed lines correspond to $0 \leq \theta \leq \pi$ and $-\pi \leq \theta \leq 0$, respectively. The innermost boundaries are the stability boundaries.

with its origin at {0,0} on the complex plane and spirals outward as the metrology delay increases. For every run that the delay is increased, the complex stability spirals outward by half a rotation around the point {1,0} and the stability boundary becomes smaller. The innermost enclosed region represents the true stability boundary and the outer rings are non-critical roots of the characteristic equation (i.e., roots which are outside of the unit circle).

Figure 2.7 illustrates geometrically how the stability boundaries tighten with increasing metrology lags. Solving for analytical solutions that describes these bounds, however, is not trivial. Since the Routh-Hurwitz criterion is only valid for polynomials with all real coefficients, the SISO stability region is valid only when all of the eigenvalues of the model mismatch matrix are real. To determine if all the poles of a polynomial with complex coefficients are in the LHP, Frank and Gantmacher independently introduced methods to extend the Routh-Hurwitz stability criterion to polynomials with complex coefficients [29, 31]. Gantmacher's method is somewhat more straightforward in its implementation so this method will be used exclusively to solve for the stability boundaries when model mismatch matrix has complex eigenvalues.

According to Gantmacher's generalized Routh-Hurwitz stability criterion, a polynomial of the form

$$f(iw) = (a_p + b_p i)w^p + (a_{p-1} + b_{p-1} i)w^{p-1} + (a_0 + b_0 i) \tag{2.33}$$

is stable when the determinants of all of the principal minors of the matrix,

$$\nabla = \begin{bmatrix} b_p & b_{p-1} & \cdots & b_o & 0 \\ a_p & a_{p-1} & \cdots & a_o & 0 \\ 0 & b_p & b_{p-1} & \cdots & b_o \\ 0 & a_p & a_{p-1} & \cdots & a_o \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \tag{2.34}$$

are greater than zero.

The generalized Routh-Hurwitz stability criterion will be demonstrated on the MIMO systems with metrology delays of one and two runs.

2.2.2.1 The Stability of a MIMO Process With a Delay of One Run

Starting with the characteristic equation of the SISO case with a delay of one run (2.11) and substituting $w = wi$ and $\lambda_j = \mathrm{a} + \mathrm{b}i$ yields the complex polynomial,

$$\begin{array}{l} \Re\{f(iw)\} = -(\mathrm{a}(1-\omega)+2\omega)w^2 + 2\mathrm{b}(1-\omega)w + \mathrm{a}(1-\omega) \\ \Im\{f(iw)\} = -\mathrm{b}(1-\omega)w^2 - 2(\mathrm{a}+\omega-\mathrm{a}\omega-2)w + \mathrm{b}(1-\omega) \end{array} \tag{2.35}$$

so that the ∇ matrix is

$$\nabla = \begin{bmatrix} b_2 & b_1 & b_0 & 0 \\ a_2 & a_1 & a_0 & 0 \\ 0 & b_2 & b_1 & b_0 \\ 0 & a_2 & a_1 & a_0 \end{bmatrix} \tag{2.36}$$

where

$$\begin{array}{cc} b_0 = \mathrm{b}(1-\omega) & a_0 = \mathrm{a}(1-\omega) \\ b_1 = -2(\mathrm{a}+\omega-\mathrm{a}\omega-2) & a_1 = 2\mathrm{b}(1-\omega) \\ b_2 = -\mathrm{b}(1-\omega) & a_2 = -\mathrm{a}(1-\omega) - 2\omega. \end{array}$$

The solution to the system of inequalities such that the principal minors of ∇ are greater than zero yields multiple stability boundaries. However, the active stability boundary results from the solution of a fourth order polynomial in x,

$$\omega = roots\left(a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\right) \tag{2.37}$$

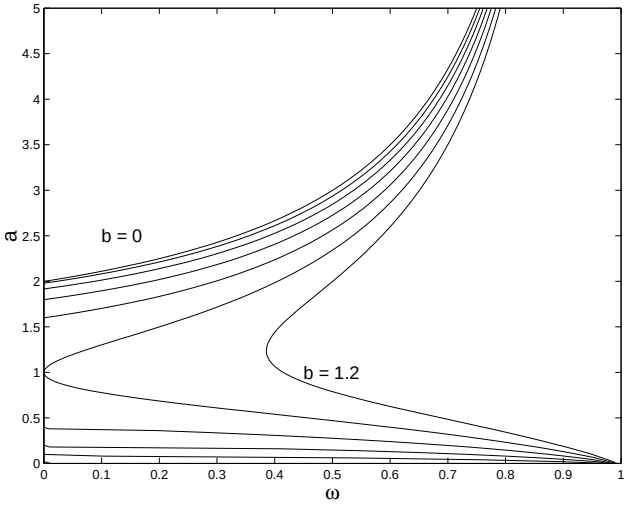

Figure 2.8: The complex stability boundaries for a system with metrology delay of one run, b = 0 . . . 1.2 in increments of 0.2

where

$$\begin{aligned}
a_0 &= 8\mathrm{b}^2\mathrm{a}^2 + 16\mathrm{a}^2 - 16\mathrm{ab}^2 - 16\mathrm{a}^3 + 4\mathrm{a}^4 + 4\mathrm{b}^4 \\
a_1 &= -16\mathrm{b}^2 - 32\mathrm{b}^2\mathrm{a}^2 + 64\mathrm{ab}^2 - 80\mathrm{a}^2 + 64\mathrm{a}^3 + 32\mathrm{a} - 16\mathrm{b}^4 - 16\mathrm{a}^4 \\
a_2 &= -96\mathrm{ab}^2 + 36\mathrm{b}^2 + 132\mathrm{a}^2 + 24\mathrm{b}^4 + 48\mathrm{b}^2\mathrm{a}^2 + 24\mathrm{a}^4 - 64\mathrm{a} - 96\mathrm{a}^3 \\
a_3 &= -32\mathrm{b}^2\mathrm{a}^2 - 16\mathrm{a}^4 + 64\mathrm{a}^3 + 40\mathrm{a} - 24\mathrm{b}^2 - 16\mathrm{b}^4 + 64\mathrm{ab}^2 - 88\mathrm{a}^2 \\
a_4 &= -16\mathrm{ab}^2 + 4\mathrm{b}^4 + 8\mathrm{b}^2\mathrm{a}^2 + 20\mathrm{a}^2 + 4\mathrm{b}^2 + 4\mathrm{a}^4 - 16\mathrm{a}^3 - 8\mathrm{a}.
\end{aligned}$$

A numerical root solver was used to find the solution to (2.37) and the solution is plotted in Figure 2.8.

2.2.2.2 The Stability of a MIMO Process With a Delay of Two Runs

Following the same procedure as the case with a delay of one run, the complex characteristic equation for a process with a delay of two runs is

$$\begin{aligned}
\Re\{f(iw)\} &= \mathrm{b}(1-\omega)w^3 + (3\mathrm{a}(1-\omega) + 4\omega)w^2 \\
&\quad -3\mathrm{b}(1-\omega)w - \mathrm{a}(1-\omega) \\
\Im\{f(iw)\} &= (2 - \mathrm{a}(1-\omega))w^3 + 3\mathrm{b}(1-\omega)w^2 \\
&\quad +(3\mathrm{a}(1-\omega) + 4\omega - 6)w - \mathrm{b}(1-\omega)
\end{aligned} \tag{2.38}$$

so that the ∇ matrix is

$$\nabla = \begin{bmatrix} b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & b_3 & b_2 & b_1 & b_0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 \end{bmatrix} \tag{2.39}$$

where

$$\begin{array}{cc} b_0 = -\mathrm{b}(1-\omega) & a_0 = -\mathrm{a}(1-\omega) \\ b_1 = 3\mathrm{a}(1-\omega) + 4\omega - 6 & a_1 = -3\mathrm{b}(1-\omega) \\ b_2 = 3\mathrm{b}(1-\omega) & a_2 = 3\mathrm{a}(1-\omega) + 4\omega \\ b_3 = 2 - \mathrm{a}(1-\omega) & a_3 = \mathrm{b}(1-\omega). \end{array}$$

The active stability boundary results from the solution of a fifth order polynomial in x,

$$\omega = roots\left(a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\right) \tag{2.40}$$

where

$$\begin{array}{c} a_0 = -\left(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a}\right)^3 \\ a_1 = (5\mathrm{a}^2 + 5\mathrm{b}^2 - 10\mathrm{a} + 6)\left(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a}\right)^2 \\ a_2 = -(10\mathrm{a}^2 + 10\mathrm{b}^2 - 20\mathrm{a} + 9)(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a} + 1)(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a}) \\ a_3 = 10\mathrm{a}^6 + 30\mathrm{a}^4\mathrm{b}^2 + 30\mathrm{a}^2\mathrm{b}^4 + 10\mathrm{b}^6 - 60\mathrm{a}^5 - 120\mathrm{a}^3\mathrm{b}^2 - 60\mathrm{ab}^4 + 141\mathrm{a}^4 \\ +162\mathrm{a}^2\mathrm{b}^2 + 21\mathrm{b}^4 - 164\mathrm{a}^3 - 84\mathrm{ab}^2 + 97\mathrm{a}^2 + 13\mathrm{b}^2 - 24\mathrm{a} \\ a_4 = -(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a})(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a} + 1)(5\mathrm{a}^2 + 5\mathrm{b}^2 - 10\mathrm{a} + 4) \\ a_5 = (\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a} + 1)\left(\mathrm{a}^2 + \mathrm{b}^2 - 2\mathrm{a}\right)^2. \end{array}$$

Solving (2.40) numerically yields the stability boundaries plotted in Figure 2.9.

2.2.2.3 The Stability of MIMO Processes with Large Metrology Delays

Like the SISO case, the procedure used to find the stable region of the system with metrology lags of one or two runs can be extended to find the stable regions for processes with longer delays. However, when metrology delays increase beyond one or two runs, the solution to the generalized Routh-Hurwitz criteria becomes

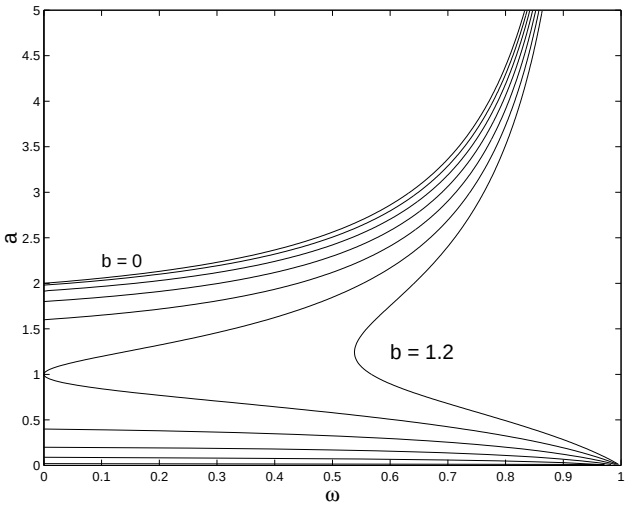

Figure 2.9: The complex stability boundaries for a system with metrology delay of two runs, $\mathrm{b} = 0 \ldots 1.2$ in increments of 0.2

prohibitively difficult. A sufficient condition for MIMO stability can be established by ensuring that the most aggressive controller, which corresponds to $\omega = 0$, is stable. The resulting characteristic equation is

$$f(z) = z^{d+1} + \lambda_j - 1 = 0. \tag{2.41}$$

By restricting the the roots of the characteristic equation to the inside of the unit circle leads to the following theorem:

Theorem 2.2.3. *A sufficient condition for stability of the MIMO EWMA controller for a delay of any length is, for* $\lambda_j = \mathrm{a} + \mathrm{b}i$*,*

$$(\mathrm{a} - 1)^2 + \mathrm{b}^2 < 1 \tag{2.42}$$

Proof: See Appendix A.4.

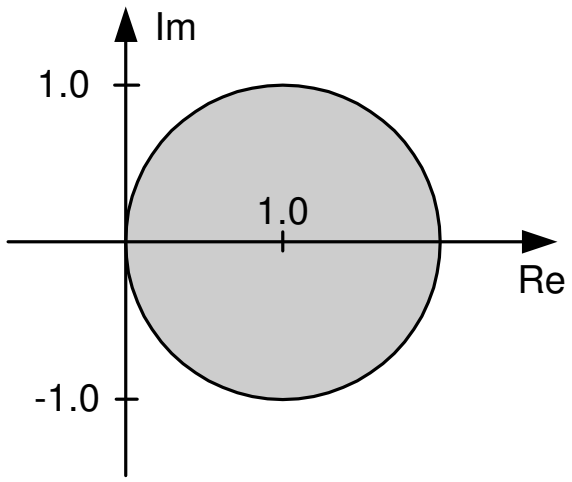

Figure 2.10: A sufficient condition for MIMO stability is that all the eigenvalues of the model-mismatch matrix fall inside a circle with unit radius and centered at {1,0} on the complex plane

This could equivalently be stated that a sufficient condition for MIMO stability is that all of the eigenvalues of the model mismatch matrix fall inside a circle centered at {1,0} and with unit radius. This is shown in Figure 2.10. By denoting $\lambda_j = \mathrm{a}+\mathrm{b}i = re^{i\theta}$, the stability boundary for $\theta = 0 \ldots 3\pi/8$ are shown in Figure 2.11 which verifies the sufficient stability condition in Theorem 2.2.3.

Since the addition of metrology delay shrinks the stability region, a necessary condition for closed-loop stability is that all of the eigenvalues of $\boldsymbol{\xi}$ meet the stability requirement for the case with no metrology delay shown in (2.29). By setting d=0 in (2.32), the stability boundary in (2.29) can be written in polar notation,

$$\lambda_j = \frac{1}{1-\omega}\left(1 - e^{i\theta}\right). \tag{2.43}$$

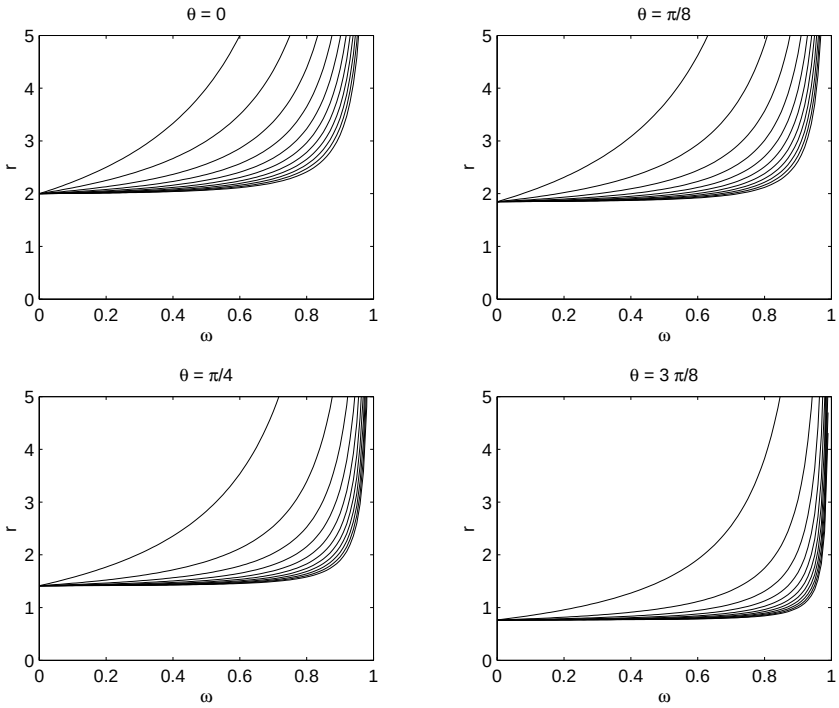

Figure 2.11: The magnitude of the stability boundary, r, and phase angle, θ, for delays of zero through ten runs

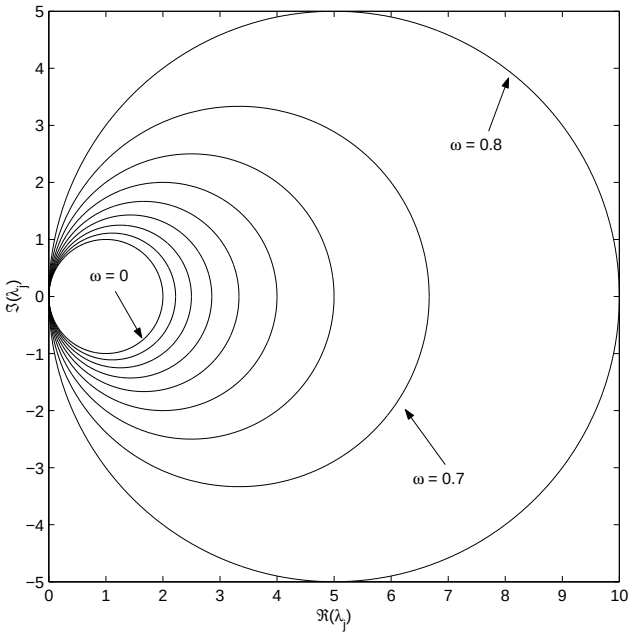

Figure 2.12: A necessary condition for stability is that all of the eigenvalues of $\boldsymbol{\xi}$ fall inside a circle with radius $r = \frac{1}{1-\omega}$ and centered at $\{\frac{1}{1-\omega}, 0\}$ on the complex plane

This states that a necessary condition for closed-loop stability for any length metrology delay is that all of the eigenvalues of the model-mismatch matrix fall inside a circle with radius, $r = \frac{1}{1-\omega}$ and centered at $\{\frac{1}{1-\omega}, 0\}$ on the complex plane. This is plotted in Figure 2.12.

2.2.3 Simulation of a MIMO Process

As an example of a process under MIMO EWMA control, consider a system with $T = \left[\begin{array}{cc} 0 & 0 \end{array}\right]^T$, $\nu = \left[\begin{array}{cc} 0 & 0 \end{array}\right]^T$,

$$\beta = \left[\begin{array}{cccc} 6 & -3 & 8 & 16 \\ 14 & 10 & -1 & 19 \end{array}\right],$$

and

$$b = \begin{bmatrix} -0.5 & -0.6 & 2.8 & 4.5 \\ 1.5 & 1.5 & 5 & 5.2 \end{bmatrix}.$$

In addition, the process has normally distributed noise of the appropriate dimension with unit variance. A MIMO EWMA controller will be used with an objective function of minimizing the 2-norm of the controller input according to (1.7) so that

$$\boldsymbol{\xi} = \begin{bmatrix} 2.0 & 1.0 \\ -2.0 & 3.6 \end{bmatrix}$$

and the eigenvalues of $\boldsymbol{\xi}$ are $\lambda = 2.8 \pm 1.2i$. The eigenvalues of $\boldsymbol{\xi}$ clearly do not meet the sufficient condition for stability (i.e., they fall outside of the circle show in Figure 2.10). However, according to the boundary regions shown in Figures 2.6 through 2.9, $\omega = 0.62$ will result in a stable system when there is no delay, a marginally stable system with a delay of one, and an unstable system with a delay of two runs. Figure 2.13 shows 50 runs of this simulation. In addition, the minimum ω required to maintain stability for various metrology delays is summarized in Table 2.1.

Table 2.1: The minimum ω required to maintain stability in the MIMO EWMA controlled example for various lengths of metrology delay

d	ω
0	0.394
1	0.623
2	0.741
3	0.805
4	0.844
5	0.870
10	0.929
∞	1.000

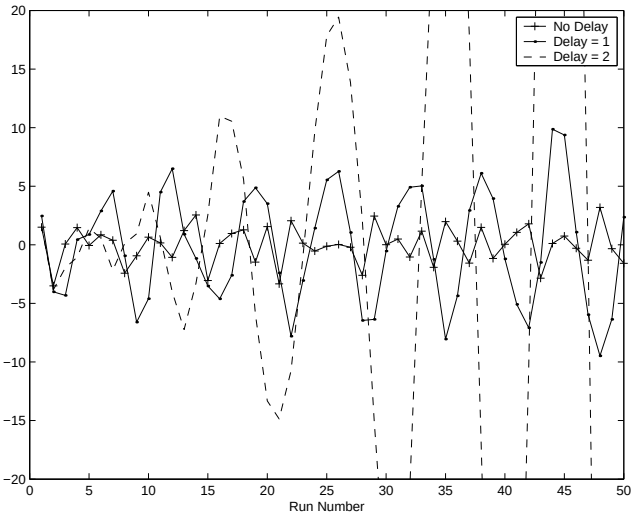

Figure 2.13: Simulation of a MIMO EWMA controller with $\lambda = 2.8 \pm 1.2i$, $\omega = 0.62$, and delays of 0, 1, and 2 runs

2.3 Conclusions

A fundamental concern in choosing the tuning parameter of the EWMA run-to-run controller is closed loop stability when there is uncertainty in the process model. This chapter has derived the tuning parameters which guarantee closed loop stability for EWMA controlled processes with time delay between product manufacturing and metrology. It was shown that for the SISO process, the sufficient condition for stability is that the true process gain must be greater than zero and less than two times the estimated process gain for any length of metrology delay. Therefore, by overestimating the process gain, the closed loop stability of a SISO system can be guaranteed.

The stability requirements for MIMO systems with several different objec-

tive functions are also derived in this chapter. The necessary and sufficient condition for MIMO stability is derived for the case of no metrology delay. For the case with metrology delay, we have shown that a sufficient condition for MIMO stability is that all of the eigenvalues of the model mismatch matrix fall inside a circle with unit radius and centered at $\{1, 0\}$ on the complex plane. Unlike the SISO case, the imaginary component of the eigenvalues can result in an unstable region when the process gain is overestimated. It follows that overestimating the process gain can result in an unstable system with MIMO control. Simulations have been shown where stable processes are pushed into the unstable region when changes in manufacturing increase the metrology lag. As a result one can conclude that it is necessary to consider both the worst case model mismatch and the metrology lag when choosing the EWMA tuning parameter for a MIMO process.

Chapter 3

The Stability of the Double-EWMA Controller

The double exponentially weighted moving average (double-EWMA) run-to-run controller is a popular algorithm for controlling semiconductor manufacturing processes which have a deterministic drift due to process aging or tool performance deterioration over time. Like the EWMA controller discussed in Chapter 2, a poorly tuned double-EWMA controller can result in closed-loop instability. This chapter examines the stability bounds for the tuning parameter of both the single input-single output (SISO) and multiple input-multiple output (MIMO) double-EWMA controllers that have model mismatch and delay between manufacturing and metrology. To illustrate the importance of performing stability analysis in choosing the double-EWMA tuning parameter, a simulation of a chemical mechanical planarization process that is controlled by a double-EWMA controller is included.

As mentioned in Section 1.1.1.2, there are three formulations of the double-EWMA controller commonly seen in the literature: the predictor corrector controller (PCC) [11], the double-EWMA controller introduced in [19] as a correction to the PCC, and the double-EWMA controller as derived from recursive least squares regression [1]. The first two approaches to the double-EWMA controller use separate tuning parameters for the slope estimate and the intercept estimate. The third, double-EWMA control algorithm, on the other hand, uses a recursive least squares

regression to fit a static model to exponentially discounted data. This interpretation of the double-EWMA controller has a single tuning parameter which is used to govern how quickly process data is discounted. Although the first two interpretations can be considered more general (because of the two degrees of freedom in the controller tuning), the choice of tuning parameters can be difficult. For this reason and for simplifying the analysis of the stability conditions, only Abraham and Ledolter's recursive least squares algorithm will be considered in this chapter. Nevertheless, the stability analysis techniques demonstrated in this chapter can easily be extended to the two-tuning-parameter double-EWMA controllers.

The outline of this chapter is as follows. The stability boundaries of the SISO double-EWMA controller is derived in Section 3.1 followed by the MIMO stability boundaries in Section 3.2. Illustrative simulations based on the CMP process are provided in Sections 3.1.3 and 3.2.3 for the SISO and MIMO controllers, respectively. Some concluding remarks are given in Section 3.3.

3.1 Stability of the SISO Double-EWMA Controller

In this section the stability boundaries are derived for the SISO double-EWMA controller. The case with no metrology delay is examined in Section 3.1.1 followed by processes with metrology delays in Section 3.1.2. The stability of processes with extended metrology delays are considered in Section 3.1.2.3. A process simulation demonstrating the utility of the double-EWMA stability analysis is then shown in Section 3.1.3.

3.1.1 SISO Stability Without Metrology Delay

By combining the updating equations for the slope and intercept estimates in (1.18) with the control law in (1.5), the closed loop realization of the double-EWMA controller can be written

$$\begin{aligned} \hat{\nu}_t &= \hat{\nu}_{t-1} + \hat{\delta}_{t-1} + (1-\omega^2)\left(-\xi(\hat{\nu}_{t-1} + \hat{\delta}_{t-1}) + \nu_{n-1} + \delta\right) \\ \hat{\delta}_t &= \hat{\delta}_{t-1} + (1-\omega)^2\left(-\xi(\hat{\nu}_{t-1} + \hat{\delta}_{t-1}) + \nu_{n-1} + \delta\right) \end{aligned} \tag{3.1}$$

where the process target, T, has been set to zero without loss of generality. Next, by defining $x_t = \begin{bmatrix} \hat{\nu}_t & \hat{\delta}_t \end{bmatrix}^T$, the updating equations in (3.1) can be written in state space representation [14],

$$\begin{aligned} x_{t+1} &= Ax_t + w_t \\ y_t &= Cx_t + R_t. \end{aligned} \tag{3.2}$$

where

$$A = \begin{bmatrix} 1-(1-\omega^2)\xi & 1-(1-\omega^2)\xi \\ -(1-\omega)^2\xi & 1-(1-\omega)^2\xi \end{bmatrix}$$

$$w_t = \begin{bmatrix} (1-\omega^2)[(\xi-1) + \nu_t + \delta + \varepsilon_t] \\ (1-\omega)^2[(\xi-1) + \nu_t + \delta + \varepsilon_t] \end{bmatrix}$$

$$C = \begin{bmatrix} -\xi \\ -\xi \end{bmatrix}$$

$$R_t = \nu_t + \varepsilon_t$$

The system described by (3.2) is stable when the both the disturbance, ν_t, and the state transition matrix, A, are bounded as t goes to infinity. It is widely known that the second condition will hold if all of the eigenvalues of A are inside of the unit circle. The eigenvalues of A are determined from the roots of the characteristic equation,

$$f(z) = z^2 + (-2\xi\omega + 2\xi - 2)z + (\omega^2\xi - \xi + 1) = 0. \tag{3.3}$$

To solve for $\xi = f(\omega)$ such that all of the eigenvalues are inside the unit circle, the inside of the unit circle is mapped to the (LHP) using the bilinear transformation in (2.10) which yields

$$f(w) = (\omega^2\xi + 2\omega\xi - 3\xi + 4)w^2 + (-2\omega^2\xi + 2\xi)w + \omega^2\xi - 2\omega\xi + \xi = 0. \quad (3.4)$$

Since (3.4) is quadratic, all the roots are in the LHP when all of the coefficients on w have the same sign. This results in the following three criteria for the stability boundaries:

1. $\omega^2\xi + 2\omega\xi - 3\xi + 4 > 0$
2. $-2\omega^2\xi + 2\xi > 0$
3. $\omega^2\xi - 2\omega\xi + \xi > 0.$

solving the second and third inequality both yield $\xi > 0$ and solving the first inequality yields

$$\xi < \frac{-4}{\omega^2 + 2\omega - 3}$$

It follows that the stability region of the SISO controller is

$$0 < \xi < \frac{-4}{\omega^2 + 2\omega - 3}. \quad (3.5)$$

This is plotted in Figure 3.1.

An equivalent derivation of the stability region shown in (3.5) can be found in [14] for Butler and Stefani's PCC interpretation of the double-EWMA controller.

As shown in Figure 3.1, a necessary condition for stability is that ξ is greater than zero or, equivalently, when the estimated process gain has the same sign as the

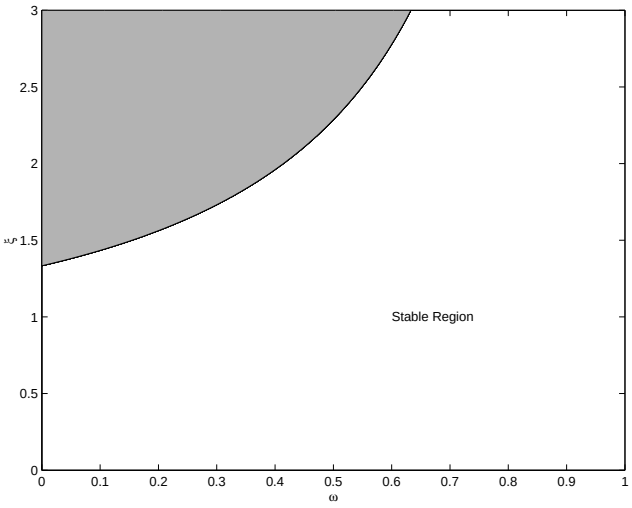

Figure 3.1: Stability region for the SISO double-EWMA controller without metrology delay

actual process gain. Figure 3.1 also demonstrates that the stability of the system increases as the forgetting factor, ω, approaches unity. This is not surprising as a larger forgetting factor is associated with a less aggressive controller. The limiting case is when $\omega = 1$. In this case, the disturbance parameter estimates, $\hat{\nu}$ and $\hat{\delta}$, are never updated and the only criterion for stability is that the disturbance is bounded. Therefore, we see that the upper stability boundary is $\xi = \infty$ when $\omega = 1$.

3.1.2 SISO Stability With Metrology Delay

When there is delay between processing and measuring a batch of wafers, the disturbance estimate must be forecasted through the duration of the metrology delay. In this case, instead of using the disturbance model to simply forecast the disturbance one step ahead (i.e., time $t + 1$), the disturbance model is used to find

the $d+1$ step ahead forecast,

$$\hat{\nu}_{t+d+1} = \hat{\nu}_t + (d+1)\hat{\delta}_t. \tag{3.6}$$

Here we see that the most recent estimate of the drift term, $\hat{\delta}_t$, is used to account for the drift during the delay. It follows that the SISO control law is

$$u_{t+1} = \frac{T - \hat{\nu}_{t-d} - (1+d)\hat{\delta}_{t-d}}{b} \tag{3.7}$$

and the parameter updating equations and the control recipe is calculated from

$$\begin{aligned} \hat{\nu}_{t-d} &= \hat{\nu}_{t-d-1} + \hat{\delta}_{t-d-1} + (1-\omega^2)\left(y_{t-d} - \hat{y}_{t-d|t-d-1}\right) \\ \hat{\delta}_{t-d} &= \hat{\delta}_{t-d-1} + (1-\omega)^2\left(y_{t-d} - \hat{y}_{t-d|t-d-1}\right) \end{aligned} \tag{3.8}$$

where

$$\hat{y}_{t-d|t-d-1} = bu_{t-d} + \hat{\nu}_{t-d-1} + \hat{\delta}_{t-d-1}. \tag{3.9}$$

To put the above system of difference equations into state space form, we define x_t to be

$$x_t = \left[\begin{array}{ccccccc} \hat{\nu}_{t-d-1} & \hat{\nu}_{t-d} & \dots & \hat{\nu}_t & \hat{\delta}_{t-d-1} & \hat{\delta}_{t-d} & \dots & \hat{\delta}_t \end{array}\right]^T. \tag{3.10}$$

The transition matrix, A, which determines the stability of the system, is then given by

$$A = \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right] \tag{3.11}$$

where

$$A_{11} = \left[\begin{array}{cccc} 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & & \ddots \\ (1-\omega^2)(1-\xi) & 0 & \dots & \omega^2 \end{array}\right]$$

$$A_{12} = \left[\begin{array}{cccc} 0 & 0 & \dots & 0 \\ 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ (d+1)(1-\omega^2)(1-\xi) & 0 & \dots & \omega^2 \end{array}\right]$$

$$A_{21} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ (1-\omega)^2(1-\xi) & 0 & \dots & -(1-\omega)^2 \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & & \ddots \\ (d+1)(1-\omega)^2(1-\xi) & 0 & \dots & \omega(2-\omega) \end{bmatrix}.$$

As with the system with no delay, the system is stable when the disturbance is bounded and all of the eigenvalues are inside the unit circle. The Routh-Hurwitz criterion can be used to solve for the boundary which ensures that all of the eigenvalues are inside the unit circle.

3.1.2.1 The Stability of a SISO System With a Delay of One Run

We first consider the stability region of a process with a delay of one run such that,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ (1-\omega^2)(1-\xi) & \omega^2 & 2(1-\omega^2)(1-\xi) & \omega^2 \\ 0 & 0 & 0 & 1 \\ (1-\omega)^2(1-\xi) & -(1-\omega)^2 & 2(1-\omega)^2(1-\xi) & 2\omega-\omega^2 \end{bmatrix}. \tag{3.12}$$

The characteristic equation of A is

$$f(z) = z^3 - 2\omega z^2 + \xi\left((1-\omega)(3-\omega) + 4\omega - 3\right)z + 2(1-\omega)(1-\xi) = 0 \tag{3.13}$$

and by using the bilinear transformation in (2.21), the inside of the unit circle in the z-domain is mapped to the LHP in the w-domain. This transformation yields

$$f(w) = a_3 w^3 + a_2 w^2 + a_1 w + a_0 = 0 \tag{3.14}$$

where

$$a_0 = -\xi(1-\omega)^2$$
$$a_1 = -\xi(3+\omega)(1-\omega)$$
$$a_2 = -\xi(1-\omega)(9-\omega) + 8\omega - 12$$
$$a_3 = \xi(1-\omega)(5+\omega) + 4(1-2\omega).$$

The Routh-Hurwitz stability criterion is then used to determine the conditions for which the roots of (3.14) are all in the LHP. This criterion states that the determinants of all of the principal minors of the Routh-Hurwitz matrix are greater than zero:

1. $a_3 > 0$
2. $\begin{vmatrix} a_3 & a_1 \\ a_2 & a_0 \end{vmatrix} > 0.$

Solving this system of inequalities yields the following stability boundaries for the system with a delay of one run:

$$\frac{-8\omega + 4}{\omega^2 - 6\omega + 5} < \xi < \frac{3\omega - 5}{4\omega - 4}. \tag{3.15}$$

This is plotted in Figure 3.2. Here we see that, unlike the case without metrology delay (where overestimating the process gain guarantees closed-loop stability), an unstable region exists in the region $0 \leq \xi \leq 1$. This implies that if a system has metrology delay then overestimating the process gain can cause the system to become unstable. We also see in Figure 3.2 that the size of the stable region becomes smaller with the inclusion of metrology delay. In the following section, we will look at the effects of further increasing the metrology delay on the stability of the double-EWMA controller.

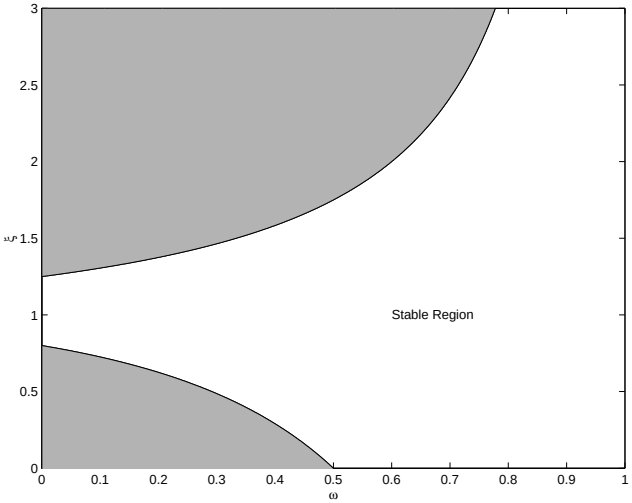

Figure 3.2: Stability region for the SISO double-EWMA controller with a delay of one run

3.1.2.2 The Stability of a SISO System With a Delay of Two Runs

According to (3.11), a system with a metrology delay of two runs has a transition matrix,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ (1-\omega)^2(1-\xi) & 0 & \omega^2 & 3(1-\omega)^2(1-\xi) & 0 & \omega^2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ (1-\omega)^2(1-\xi) & 0 & -(1-\omega)^2 & 3(1-\omega)^2(1-\xi) & 0 & \omega(2-\omega) \end{bmatrix}$$

with a characteristic polynomial

$$\begin{aligned} f(z) &= z^4 - 2\omega z^3 + \omega^2 z^2 + 2(\omega-2)(1-\omega)(1-\xi)z \\ &\quad -(\omega-3)(1-\omega)(1-\xi) = 0. \end{aligned} \tag{3.16}$$

Applying the bilinear transformation to (3.16) yields the characteristic polynomial in w,

$$f(w) = a_4w^3 + a_3w^3 + a_2w^2 + a_1w + a_0 = 0 \tag{3.17}$$

where

$$\begin{gathered} a_0 = \xi(1-\omega)^2 \\ a_1 = 4\xi(1-\omega) \\ a_2 = (-6\xi+4)\omega^2 + (24\xi-24)\omega - 18\xi + 24 \\ a_3 = (8\xi-8)\omega^2 + (-28\xi+32)\omega + 20\xi - 16 \\ a_4 = (-3\xi+4)\omega^2 + (10\xi-8)\omega - 7\xi + 8. \end{gathered}$$

It follows that the Routh-Hurwitz criterion results in the following requirements for stability:

1. $a_4 > 0$
2. $a_3 > 0$
3. $\begin{vmatrix} a_3 & a_1 \\ a_4 & a_2 \end{vmatrix} > 0$
4. $\begin{vmatrix} a_3 & a_1 & 0 \\ a_4 & a_2 & a_0 \\ 0 & a_3 & a_1 \end{vmatrix} > 0.$

Solving this system of inequalities results in the following possible stability boundaries:

1. $\xi < \frac{4(\omega^2-2\omega+2)}{3\omega^2-10\omega+7}$
2. $\xi > \frac{2(\omega^2-4\omega+2)}{2\omega^2-7\omega+5}$
3. (a) $\xi > \frac{10\omega^3-74\omega^2+165\omega-92+2\sqrt{\omega^6-4\omega^5-5\omega^4+28\omega^3-8\omega^2-56\omega+124}}{(-12\omega^2+63\omega-83)(1-\omega)}$,

   (b) $\xi < \frac{10\omega^3-74\omega^2+165\omega-92-2\sqrt{\omega^6-4\omega^5-5\omega^4+28\omega^3-8\omega^2-56\omega+124}}{(-12\omega^2+63\omega-83)(1-\omega)}$

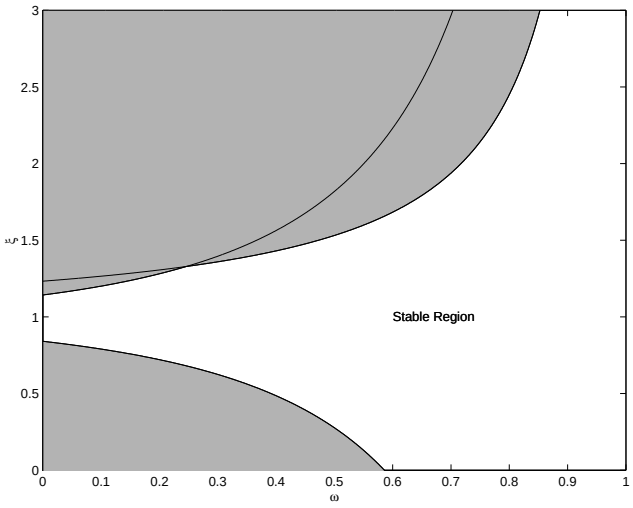

Figure 3.3: Stability region for the SISO double-EWMA controller with a delay of two runs

4. (a) $\xi > 0$,

   (b) $\xi > \frac{\omega^4-10\omega^3+37\omega^2-58\omega+28+2\sqrt{-2\omega+7}}{(\omega^3-9\omega^2+27\omega-27)(\omega-1)}$,

   (c) $\xi < \frac{\omega^4-10\omega^3+37\omega^2-58\omega+28-2\sqrt{-2\omega+7}}{(\omega^3-9\omega^2+27\omega-27)(\omega-1)}$.

The active stability boundaries are determined by evaluating and comparing each of the above choices. It follows that the active stability boundaries for the double-EWMA controller with a delay of two runs are

$$\frac{\omega^4-10\omega^3+37\omega^2-58\omega+28+2\sqrt{-2\omega+7}}{(\omega^3-9\omega^2+27\omega-27)(\omega-1)} < \xi < \left\{ \begin{array}{ll} \frac{4\left(\omega^2-2\omega+2\right)}{3\omega^2-10\omega+7} & ,\omega \leq 0.246 \\ \frac{\omega^4-10\omega^3+37\omega^2-58\omega+28-2\sqrt{-2\omega+7}}{(\omega^3-9\omega^2+27\omega-27)(\omega-1)} & ,\omega > 0.246 \end{array} \right\} \tag{3.18}$$

This region is shown in Figure 3.3.

Equation (3.18) and Figure 3.3 show that the upper active stability boundary changes at $\omega = 0.246$. When the metrology delay increases, roots are added to the

characteristic polynomial and these additional roots can become active boundaries. It can be observed empirically that each additional metrology lag adds an additional stability boundary.

3.1.2.3 A Sufficient Condition for SISO Stability

We can continue to derive the stability regions for systems with longer delays than two runs by looking at the characteristic equation,

$$f(z) = z^{d+2} + a_{d+1}z^{d+1} + a_d z^d + a_1 z + a_0 = 0 \tag{3.19}$$

where

$$\begin{gathered} a_0 = (d(1-\omega)^2 + (1+\omega)(1-\omega))(1-\xi) \\ a_1 = -(d(1-\omega)^2 + 2(1-\omega))(1-\xi) \\ a_d = \omega^2 \\ a_{d+1} = -2\omega. \end{gathered}$$

However, just as in the EWMA stability analysis in Section 2.1.2.3, a sufficient condition for stability can be derived by looking at the case when $\omega = 0$. The idea here is that by looking at the most aggressive (and therefore the most unstable) controller tuning, a sufficient condition for closed-loop stability can be inferred for less aggressive controller tunings. With this in mind, the characteristic equation with $\omega = 0$ is

$$f(z) = z^{d+2} - (d+2)(1-\xi)z + (d+1)(1-\xi) = 0. \tag{3.20}$$

Solving (3.20) for ξ yields the stability boundary

$$\xi = 1 - \frac{z^{d+2}}{z(d+2) - (d+1)} \tag{3.21}$$

where z can be any value on the unit circle that yields a real value of ξ. The right half of (3.21) is of particular interest as it represents the distance from $\xi = 1$ within

which the system remains stable. This gain margin (GM) is minimized for $z = -1$ so that the minimum distance from $\xi = 1$ to the stability boundary is

$$GM = \frac{1}{2d+3}. \tag{3.22}$$

By returning to Figures 3.1 through 3.3, we can see the utility of this sufficient condition. First, Figure 3.1 shows that the closest that the stability boundary comes to $\xi = 1$ occurs when $\xi = \frac{4}{3}$ so that the gain margin is $\frac{1}{3}$. We also see from Figures 3.2 and 3.3 that the closest distance to the stability boundary is $\frac{1}{5}$ and $\frac{1}{9}$ for delays of one and two runs, respectively. We can also infer from comparing Figures 3.1 through 3.3 that the smallest distance to the stability boundary alternates between being above and below $\xi = 1$. When the metrology delay of a process is an even number then the smallest distance to the stability boundary will occur above $\xi = 1$. Conversely, when length of the metrology delay is an odd number, then the shortest distance to the stability boundary will be below $\xi = 1$. Figure 3.4 shows a plot of the sufficient condition for stability along with the stability boundary calculated at $\omega = 0$.

3.1.2.4 A Necessary Condition for SISO Stability

For systems with delays longer than one run, a necessary condition for stability is that the magnitude of the a_0 term in (3.19) is less than one,

$$\left|(d(1-\omega)^2 + (1+\omega)(1-\omega))(1-\xi)\right| < 1. \tag{3.23}$$

Solving (3.23) for ξ yields the necessary condition for stability,

$$1 - \frac{1}{d(1-\omega)^2 + (1+\omega)(1-\omega)} < \xi < 1 + \frac{1}{d(1-\omega)^2 + (1+\omega)(1-\omega)}. \tag{3.24}$$

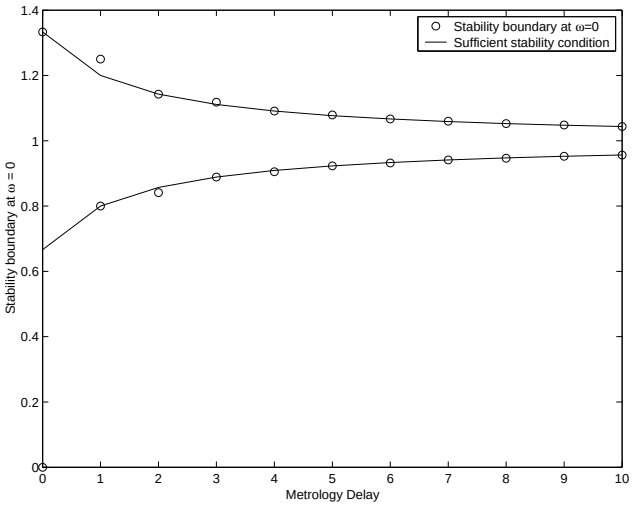

Figure 3.4: A sufficient condition for stability and the stability boundary at $\omega = 0$

This is plotted alongside the numerically determined necessary and sufficient stability conditions in Figure 3.5 for delays of two through 10 runs.

One of the key drawbacks to the double-EWMA controller (as compared to the EWMA controller) is illustrated in Figure 3.5. As the metrology delay approaches infinity, any value of $\xi \neq 1$ will be outside of the necessary condition for stability, that is,

$$\lim_{d\to\infty} 1 \pm \frac{1}{d(1-\omega)^2 + (1+\omega)(1-\omega)} = 1, \tag{3.25}$$

so that the gain margin approaches zero as the metrology delay approaches infinity. This is in contrast to the EWMA controller where the gain margin is unity, irrespective of the length of the metrology delay.

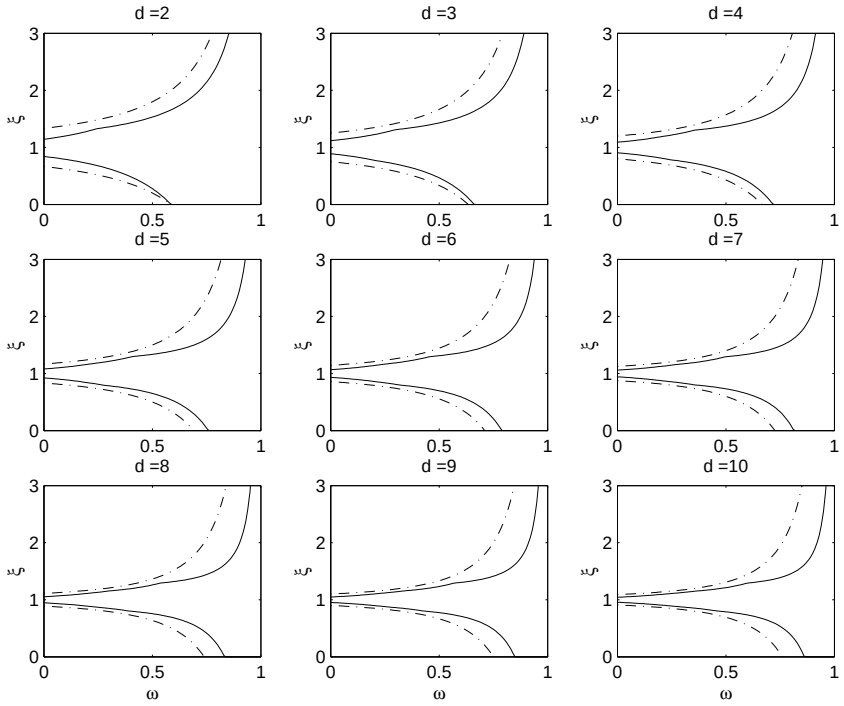

Figure 3.5: A necessary condition for stability is plotted along with the necessary and sufficient condition for stability for delays of two through ten runs. The necessary condition is plotted with a dashed line.

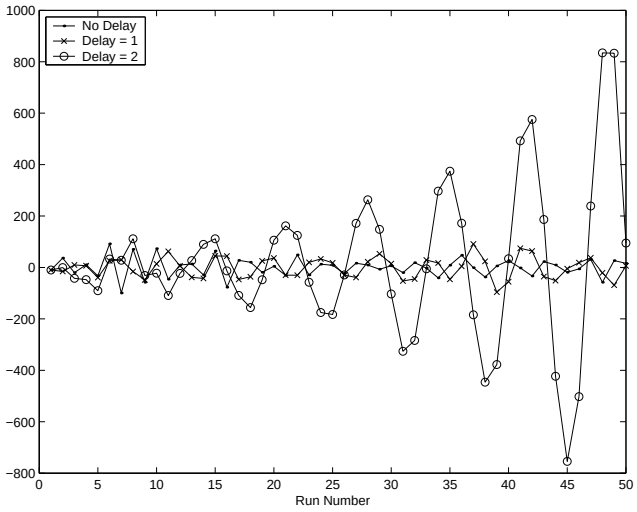

Figure 3.6: Simulation of a double-EWMA controlled process with $\omega = 0.6$ and delays of 0, 1, and 2 runs

3.1.3 Simulation of a SISO Process

To demonstrate the importance of considering delay in choosing the double-EWMA tuning parameter, consider the following example. We wish to control a system with $T = 0, \nu_0 = 0, \beta = 5, \delta = -15$, and with normally distributed noise with a standard deviation of 14.5. Through a DOE or from process data, an approximation for the model gain is determined to be $b = 2.5$ corresponding to $\xi = 2.0$. We see immediately that ξ is outside of the sufficient condition for stability so that the necessary and sufficient conditions must be used. According to the boundary regions shown in Figures 3.1 through 3.3, $\omega = 0.6$ will result in a stable system when there is no delay, a marginally stable system with a delay of one and an unstable system with a delay of two. Figure 3.6 shows 50 runs of this simulation.

It is clear from the simulation shown in Figure 3.6 that it is important to consider the effects of metrology delay on the closed loop stability of the process. This is especially important after tool maintenance events (where the process gain can change [37]) or in increasing the process throughput (where delay at the metrology tool can increase). With increasing runs, the system with no delay remains on target while the systems with delay become unstable. The scenario with a delay of one is marginally stable according to the eigenvalues of the A matrix (suggesting that oscillations should exist but not increase over time) but Figure 3.6 shows that a minor increase in the amplitude of the oscillations over time may be occurring. This is simply the result of computational error in the simulation. By increasing ω to 0.8 the eigenvalues of all three systems are inside the unit circle and, therefore, all three systems are stable.

3.2 MIMO Double-EWMA Stability With Metrology Delay

In this section the stability conditions of the MIMO double-EWMA controller are derived. The stability boundaries for a system without metrology delay is analyzed in Section 3.2.1 followed by an analysis of systems with metrology delay in Section 3.2.2. In Section 3.2.2.3 a sufficient condition for the stability of the MIMO double-EWMA system is derived. Finally, a simulation of a MIMO CMP controller is provided in Section 3.2.3.

3.2.1 MIMO Double-EWMA Stability Without Metrology Delay

As mentioned previously in Section 1.1.1.2, the double-EWMA control law is equivalent to the EWMA control law except the disturbance prediction includes a slope term. It follows that the model mismatch matrices, $\boldsymbol{\xi}$, are equivalent to those used in the EWMA stability analysis, (2.22) through (2.25). The closed-loop realization can then be written in state-space form where the transition matrix, A is given by:

$$A = \begin{bmatrix} I_m - (I_m - \boldsymbol{\omega}^2)\boldsymbol{\xi} & I_m - (I_m - \boldsymbol{\omega}^2)\boldsymbol{\xi} \\ (I_m - \boldsymbol{\omega})^2\boldsymbol{\xi} & I_m - (I_m - \boldsymbol{\omega})^2\boldsymbol{\xi} \end{bmatrix} \tag{3.26}$$

where $x_t = \begin{bmatrix} \hat{\nu}_t & \hat{\delta}_t \end{bmatrix}$ and $\boldsymbol{\omega}$ is the same forgetting factor as the SISO case multiplied by the m-dimensional identity matrix,

$$\boldsymbol{\omega} = \omega I_m = \begin{bmatrix} \omega & & \\ & \ddots & \\ & & \omega \end{bmatrix}. \tag{3.27}$$

As with the SISO case, the system is stable when all of the eigenvalues of A are inside the unit circle and the process disturbances are bounded. Finding the eigenvalues of the transition matrix can be simplified using an approach analogous to Lemma 2.2.1.

Lemma 3.2.1. *The model mismatch matrix can be replaced with its eigenvalues in determining the stability boundary of the MIMO system for $\boldsymbol{\omega} = \omega I$. That is,*

$$eig\left(A(\xi)\right) = eig\left(A(D)\right) \tag{3.28}$$

where D is a matrix with the eigenvalues of the model mismatch matrix on the diagonal

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}.$$

Proof: See Appendix B.1.

For example, according to Lemma 3.2.1, a system without metrology delay will be stable when all of the eigenvalues of

$$eig(A) = \begin{bmatrix} 1-(1-\omega^2)\lambda_j & 1-(1-\omega^2)\lambda_j \\ -(1-\omega)^2\lambda_j & 1-(1-\omega)^2\lambda_j \end{bmatrix} \qquad \forall j = 1 \ldots m \tag{3.29}$$

are inside the unit circle, where $\lambda_1 \ldots \lambda_m$ are the eigenvalues of the model mismatch matrix, $\boldsymbol{\xi}$. By comparing (3.29) with (3.2), it is easy to see that the transition matrix of the MIMO double-EWMA controller is equivalent to m SISO transition matrices with the SISO plant-model mismatch, ξ, replaced with the eigenvalues of the MIMO plant-model mismatch matrix, λ_j. Therefore, the MIMO system is stable when all of the eigenvalues of $\boldsymbol{\xi}$ are inside the stability region derived for the SISO case. This result is stated in the following theorem and is a direct result of Lemma 3.2.1.

Theorem 3.2.2. *The MIMO stability region with respect to the eigenvalues of* $\boldsymbol{\xi}$ *are equivalent to the SISO stability region with respect to* ξ*.*

Proof: See Appendix B.2.

Theorem 3.2.2 is a powerful tool in reducing the complexity in determining the stability boundaries of the double-EWMA controller. For example, the transition matrix of a three-output process ($m = 3$) would have the dimensions of 6x6. Solving this system directly (i.e., without using Theorem 3.2.2) would require solving a 6^{th} order polynomial – a nearly impossible task. According to Theorem 3.2.2; however, this system can be reduced to three independent transition matrices each with a dimension of 2x2. The solution of which is equivalent to (3.5) with respect to the eigenvalues of the model-mismatch matrix.

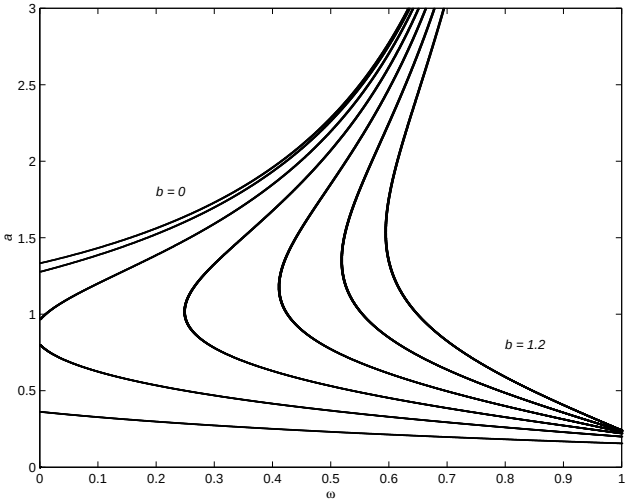

Figure 3.7: The complex stability boundaries for a system without metrology delay, b = 0 . . . 1.2 in increments of 0.2

Although a complex model mismatch parameter is not physically possible for the SISO case, the eigenvalues of $\boldsymbol{\xi}$ can be complex. Therefore, substituting $\boldsymbol{\xi} = \mathrm{a} + \mathrm{b}i$ into (3.5) and solving for ω, the stability region of the system without metrology delay can be written

$$\omega > \Re\left\{-1 \pm 2\frac{\sqrt{(\mathrm{a}+\mathrm{b}i)^2 - (\mathrm{a}+\mathrm{b}i)}}{(\mathrm{a}+\mathrm{b}i)}\right\} \tag{3.30}$$

and after multiplying through by the complex conjugate,

$$\omega > -1 + \frac{\sqrt{2(\mathrm{a}^2+\mathrm{b}^2)\left(\mathrm{a}^2 - \mathrm{a} + \sqrt{((\mathrm{a}-1)^2+\mathrm{b}^2)(\mathrm{a}^2+\mathrm{b}^2)+\mathrm{b}^2}\right)}}{a^2+b^2}. \tag{3.31}$$

The complex stability region for the double-EWMA controller without metrology delay is shown in Figure 3.7.

As shown in Figure 3.7, the case with b = 0 results in the same stability

region as shown in Figures 3.1 and an increase in the complex part of the model mismatch decreases the size of the stable region.

It is worth noting that the results in this section contradict the MIMO double-EWMA stability region published by Del Castillo and Rajagopol [15, 17]. Their results claim that the off diagonal terms of the plant mismatch matrix, $\boldsymbol{\xi}$, have no effect on the stability of the system. This, however, is only the case when the off diagonal terms have no effect on the *eigenvalues* of $\boldsymbol{\xi}$ (as is the case with diagonal, upper triangular, or lower triangular matrices) and is not true in a more general sense. A counter example to their results as well as a simple correction are included in Appendix B.3.

3.2.2 MIMO Double-EWMA Stability With Metrology Delay

Unlike the case with no metrology delay, the stability regions for the double-EWMA controller with delay is not as simple as just substituting $\boldsymbol{\xi} = \mathrm{a} + \mathrm{b}i$ into the SISO stability boundary solutions. This is not because Theorem 3.2.2 is invalid for systems with metrology delay but rather it is because the Routh-Hurwitz stability criterion is valid only for polynomials with all real coefficients. As mentioned in Section 2.2.2, the generalized Routh-Hurwitz criterion must be used when a polynomial has complex coefficients.

3.2.2.1 The stability of a MIMO System With a Delay of One Run

Starting with the characteristic equation in (3.14) and substituting $\xi = \mathrm{a}+\mathrm{b}i$ and $w = wi$ yields the complex polynomial

$$\begin{aligned} \Re\{f(iw)\} &= a_4w^4 + a_3w^3 + a_2w^2 + a_1w + a_0 = 0 \\ \Im\{f(iw)\} &= b_4w^4 + b_3w^3 + b_2w^2 + b_1w + b_0 = 0 \end{aligned} \tag{3.32}$$

where

$$\begin{aligned} b_0 &= -\omega^2\mathrm{b} - 2\omega\mathrm{a} + \mathrm{b} & a_0 &= -\omega^2\mathrm{a} + 2\omega\mathrm{b} + \mathrm{a} \\ b_1 &= -\omega\mathrm{a} + 4\mathrm{b} & a_1 &= \omega\mathrm{b} + 4\mathrm{b} \\ b_2 &= 2\omega^2\mathrm{b} + (8\mathrm{a} - 8)\omega - 6\mathrm{b} & a_2 &= 2\omega^2\mathrm{a} - 8\omega\mathrm{b} - 6\mathrm{a} + 12 \\ b_3 &= 4\omega\mathrm{a} - 4\mathrm{b} & a_3 &= -4\omega\mathrm{b} - 4\mathrm{a} + 8 \\ b_4 &= -\omega^2\mathrm{b} + (8 - 6\mathrm{a})\omega + 5\mathrm{b} & a_4 &= -\omega^2\mathrm{a} + 6\omega\mathrm{b} + 5\mathrm{a} - 4 \end{aligned}$$

so that the ∇ matrix is

$$\nabla = \begin{bmatrix} b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 & 0 \\ a_4 & a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 \\ 0 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ 0 & a_4 & a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & 0 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & a_4 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & 0 & b_4 & b_3 & b_2 & b_1 & b_0 \\ 0 & 0 & 0 & a_4 & a_3 & a_2 & a_1 & a_0 \end{bmatrix} \tag{3.33}$$

It is unreasonable to solve all of the principal minors of (3.33) analytically so a numerical solver was used. The results are plotted in Figure 3.8.

3.2.2.2 The stability of a MIMO System With a Delay of Two Runs

The same procedure as Section 3.2.2.1 can be used to find the stability boundaries of the MIMO double-EWMA process with two runs. The resulting stability boundaries are shown in Figure 3.9.

By comparing Figure 3.7 through 3.9 a few general remarks can be made concerning the stability of MIMO systems. First, as mentioned earlier, the stability

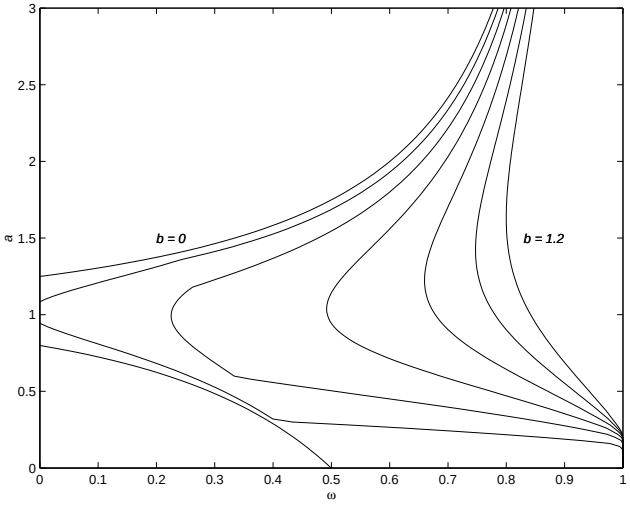

Figure 3.8: The complex stability boundaries for a system with a metrology delay of one run, b = 0 . . . 1.2 in increments of 0.2

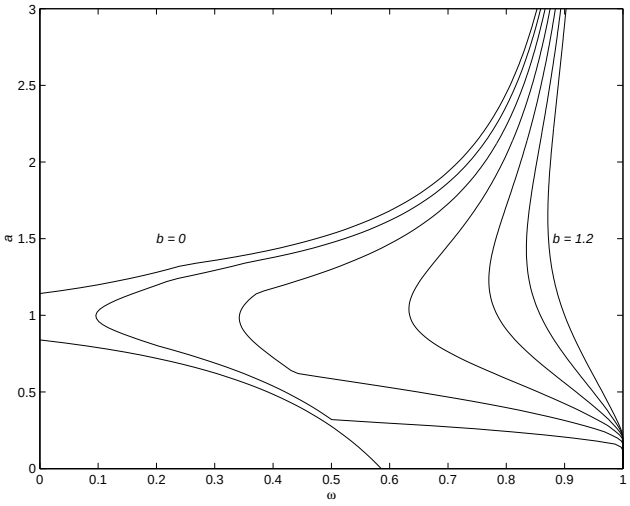

Figure 3.9: The complex stability boundaries for a system with a metrology delay of two runs, b = 0 . . . 1.2 in increments of 0.2

boundary with b = 0 is equivalent to the SISO stability boundaries in Figures 3.1 through 3.3, so, like all systems, the SISO case is simply a subset of the more general MIMO stability boundaries. Next, we see that as the complex part of the eigenvalues of the model mismatch matrix increases, the stability region is pushed to the right. When the eigenvalues are completely imaginary (i.e., a = 0) then the only stable tuning parameter is $\omega = 1$. This scenario occurs when the gains of a non-interacting system are swapped. That is, the gain for process p is assigned to process q and vice versa. It is intuitive that the only way to achieve closed-loop stability in such a system is to turn off the controller ($\omega = 1$).

3.2.2.3 A Sufficient Condition for MIMO Stability

A sufficient condition for MIMO stability can be derived by using the same approach as the one used to find a SISO sufficient stability condition in Section 3.1.2.3. By looking at the case where $\omega = 0$, the worst case (i.e., the smallest) stability boundary boundary can be derived. By setting $\omega = 0$, the transition matrix for the MIMO process is

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{bmatrix} \tag{3.34}$$

where

$$A_j = \begin{bmatrix} 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ (1-\lambda_j) & 0 & \dots & 0 & (d+1)(1-\lambda_j) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & & 0 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \ddots \\ (1-\lambda_j) & 0 & \dots & -1 & (d+1)(1-\lambda_j) & 0 & \dots & 0 \end{bmatrix}.$$

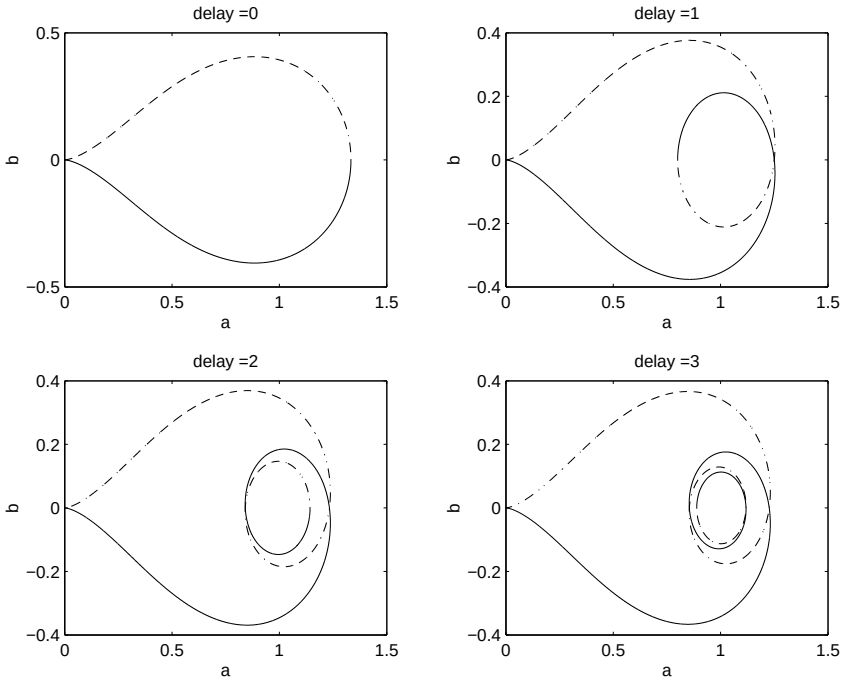

Figure 3.10: The complex stability boundary of the MIMO double-EWMA controller with $\omega = 0$ is a spiral which converges on {1,0}

From Theorem 3.2.2 and the SISO characteristic equation, (3.20), the stability boundary for each of the m disturbance channels is

$$\lambda_j = \mathrm{a}_j \pm \mathrm{b}_j i = 1 - \frac{z^{d+2}}{z(d+2) - (d+1)} \tag{3.35}$$

where z can be any point along the unit circle $z = e^{i\theta}$. By plotting (3.35) for $-\pi \leq \theta \leq \pi$, we can see that the stability boundaries are defined by a spiral in the complex plain. This is shown in Figure 3.10 for delays of zero through three runs.

The terminal point on the spiral occurs when $\theta = \pm\pi$ so that $z = -1$. It follows that the GM of the MIMO system with respect to the eigenvalues of the model mismatch matrix is equivalent to the SISO GM in (3.22) with respect to ξ.

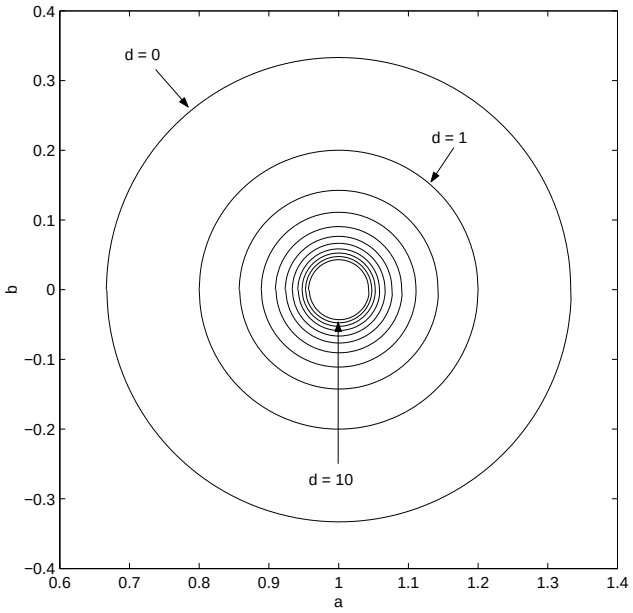

Figure 3.11: A sufficient stability condition for the double-EWMA controller for delays of zero through ten runs

This leads to the following theorem,

Theorem 3.2.3. *A sufficient condition for stability of the MIMO double-EWMA controller for a delay of length d, and for* $\lambda_j = \mathrm{a} + \mathrm{b}i$ *is*

$$(\mathrm{a}-1)^2 + \mathrm{b}^2 < \frac{1}{2d+3}. \tag{3.36}$$

Proof: See Appendix B.4.

Theorem 3.2.3 states that a sufficient condition for MIMO stability is that all of the eigenvalues of $\boldsymbol{\xi}$ fall inside a circle of radius $\frac{1}{2d+3}$ centered at {1,0} on the complex plain. This is illustrated in Figure 3.11 for delays of up to ten runs.

Figure 3.11 shows that as the metrology delay increases, the sufficient condition for stability approaches a circle centered at {1,0} on the complex plane and with zero radius. This means that as the metrology delay approaches infinity and with $\omega = 0$, any model mismatch will will be outside of this sufficient condition for stability.

3.2.2.4 A Necessary Condition for MIMO Stability

A necessary condition for MIMO stability can be derived by solving (3.23) with respect to the eigenvalues of the model-mismatch matrix,

$$\left|(d(1-\omega)^2 + (1+\omega)(1-\omega))(1-\lambda_j)\right| < 1. \tag{3.37}$$

Rewriting (3.37) in polar coordinates yields the boundary for the necessary condition for stability,

$$\lambda_j = 1 + \frac{e^{i\theta}}{d(1-\omega)^2 + (1+\omega)(1-\omega)}, \tag{3.38}$$

where $-\pi \leq \theta \leq \pi$. This is simply a circle centered at $\{1, 0\}$ in the complex plane with radius, $r = \frac{1}{d(1-\omega)^2+(1+\omega)(1-\omega)}$. A necessary condition for stability is also that the real component of λ_j is greater than zero. Combining these two necessary conditions results in the boundaries plotted in Figure 3.12.

We see in Figure 3.12 that the necessary condition for stability becomes smaller with increasing metrology delay. By taking the limit of (3.38 as d approaches infinity,

$$\lim_{d\to\infty} \lambda_j = \lim_{d\to\infty} \left\{1 + \frac{e^{i\theta}}{d(1-\omega)^2 + (1+\omega)(1-\omega)}\right\} = 1, \tag{3.39}$$

we see that the system is stable if and only if all of the eigenvalues of the model-mismatch matrix are unity.

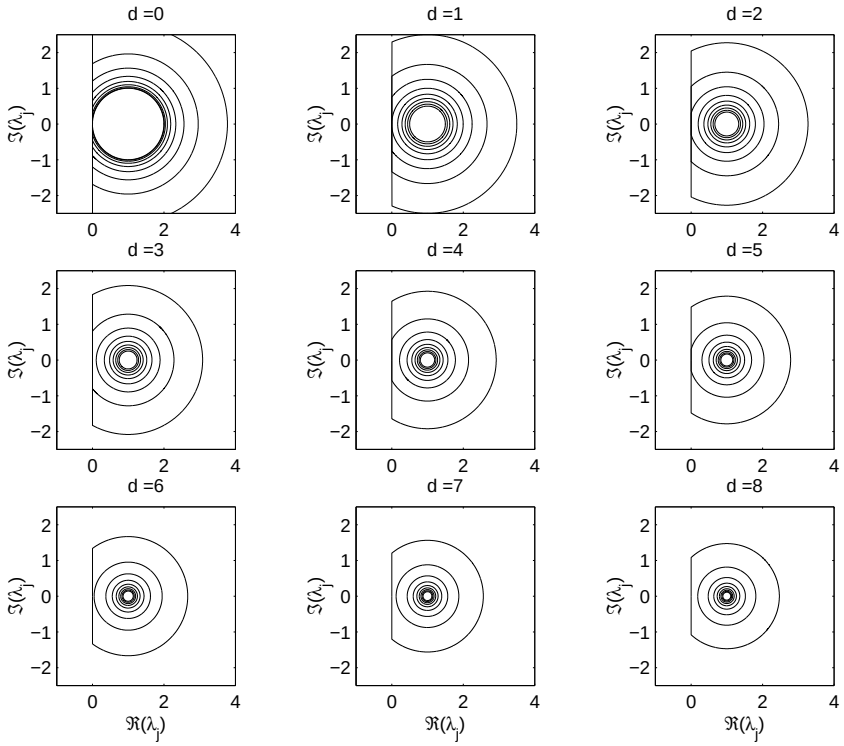

Figure 3.12: A necessary condition for stability for metrology delay of zero through eight runs for $\omega = 0 \ldots 0.8$ in increments of 0.1. The innermost circle is the $\omega = 0$ necessary condition.

3.2.3 Simulation of a MIMO Process

As a practical example, consider the chemical mechanical planarization (CMP) model described by Ning *et al* [48]. This is a four input (platen speed, back pressure, polish head down force, and profile) and two-output (removal rate and non-uniformity) system with a deterministic drift. This system matches the model given in (1.12) where ε_t, ν_t, and δ are 2x1, u is 4x1, and b is 2x4. Ning *et al* use the values

$$\nu_0 = \begin{bmatrix} -1382.60 \\ -627.32 \end{bmatrix}$$
$$\beta = \begin{bmatrix} 50.18 & -6.65 & 163.4 & 8.45 \\ 13.67 & 19.95 & 27.52 & 5.25 \end{bmatrix}$$
$$\delta = \begin{bmatrix} -17 \\ 1.5 \end{bmatrix}$$

as the parameters in their model and assume that ε_t is normally distributed and uncorrelated noise with covariance

$$\Lambda = \begin{bmatrix} 665.64 & 0 \\ 0 & 5.29 \end{bmatrix}.$$

Now, assuming that the estimate for the process gain is determined *a priori* to be

$$b = \begin{bmatrix} 19.76 & -10.96 & 70.84 & 2.20 \\ 16.59 & 11.98 & 42.97 & 4.71 \end{bmatrix},$$

it is desired to determine the values of the discount coefficient, ω, that will ensure that the system remains stable when subject to a process drift and metrology delay. We have chosen minimizing the 2-norm of the change in recipe as the objective function so that the controller input is given in (1.7). According to (2.23), the model mismatch matrix that corresponds to this objective function is

$$\boldsymbol{\xi} = \beta b^T (b b^T)^{-1} = \begin{bmatrix} 1.7 & 1.0 \\ -0.4 & 1.3 \end{bmatrix}$$

and the eigenvalues of $\boldsymbol{\xi}$ are $\lambda = 1.5 \pm 0.6i$.

We see immediately that the eigenvalues are outside of the sufficient condition for stability in (3.36) so that the full analytical solutions must be used. From the stability regions shown in Figures 3.7 through 3.9, the process under consideration will be stable when $0.348 < \omega < 1$ for the process with no delay, $0.584 < \omega < 1$ for a process with a delay of one and $0.709 < \omega < 1$ for a process with a delay of two. Table 3.1 shows the stability bounds for metrology lags of up to 20 runs.

Table 3.1: The minimum ω required to maintain stability in the CMP example for varying lengths of metrology delay

Delay	ω
0	0.348
1	0.584
2	0.709
3	0.779
5	0.851
10	0.918
20	0.957
∞	1.000

As expected, Table 3.1 shows that as metrology lag increases, the minimum value of ω that guarantees closed loop stability increases or, equivalently, the range of values of ω that guarantee closed loop stability decreases with increasing metrology delay. This demonstrates the importance in considering both the delay and the model mismatch in determining the double-EWMA tuning parameter. A well tuned, robust double-EWMA controller could be pushed into an unstable operating region if an increase in production results in an increase in the metrology lag. For example,

a tuning parameter of $\omega = 0.7$ would be stable for metrology delays of zero or one run. Then, by increasing the delay to two runs, the system becomes unstable and continues to be further outside of the stability region as the metrology lag increases.

3.3 Conclusions

Because of increased production in semiconductor manufacturing, the number of lots run between wafer processing at a tool and collecting measurements for a run continues to increase. This metrology lag can lead to process instabilities and should be considered implicitly when choosing the tuning parameters for a double-EWMA controller. This chapter has introduced a simple technique for determining the closed-loop stability region for both SISO and MIMO systems when model mismatch and measurement delay are present in a system. As expected, the size of the stability region of a SISO system decreases as metrology delay and model-mismatch increase and as the weighting parameter, ω, approaches zero. It was also shown that, in contrast to the EWMA controller, when metrology delay exists in the system, overestimating β can lead to closed-loop instability.

For a MIMO system with all real eigenvalues of the model mismatch matrix, it was shown that the necessary and sufficient condition for closed loop stability is that all of the eigenvalues must fall within the SISO stability bounds (with respect to ξ). Unlike the SISO case, however, the MIMO system can have complex eigenvalues so that the complex solution to the SISO stability region must be considered. Using the generalized Routh-Hurwitz criterion, the stability boundaries for systems with delays of up to two runs are shown. A simulation of a MIMO CMP process was

included to show the importance of considering both plant model mismatch and metrology delay in determining a robust double-EWMA tuning parameter. By increasing the metrology lag in the process the example shows that a stable system with a metrology lag of zero or one runs could be pushed into the unstable region by increasing the lag to two or more runs. To ensure that the process remains stable, a less aggressive ω should be used as metrology lag increases.

Finally, a necessary condition and a sufficient condition for stability of the SISO and MIMO double-EWMA controllers have been derived. The sufficient condition for SISO stability states that the model mismatch, ξ, must fall between a pair of bounds centered around $\xi = 1$ where the bounds decrease in magnitude as the metrology delay increases. The sufficient condition for MIMO stability is similar in that the eigenvalues of the model mismatch matrix must fall inside a circle centered at {1,0} on the complex plane. The radius of the circle decreases with metrology delay in an analogous fashion to the SISO stability boundaries. The necessary condition for stability states that as the metrology delay approaches infinity, the closed-loop system will be unstable if any of the eigenvalues are not unity. This is in contrast to the EWMA controller where the system is closed-loop stable when all of the eigenvalues are inside a circle with unity radius and centered at {1,0} on the complex plane.

Chapter 4

Approximate Solutions to the Stability Boundary Problem

In Chapters 2 and 3, analytical expressions were derived for the closed-loop stability boundaries of the EWMA and double-EWMA controllers. The equations that describe these boundaries are very simple when there is no delay between processing and metrology. However, with the introduction of metrology delay, finding analytical solutions to the stability boundaries becomes prohibitively complicated. To simplify the equations describing the stability boundaries, an accurate numerical approximation technique is introduced in this chapter for determining the stability boundaries of the EWMA and double-EWMA controllers.

The outline of this chapter is as follows. An approximation for the EWMA stability boundaries is shown in Section 4.1 followed the double-EWMA stability boundaries in Section 4.2. Some conclusions are given in Section 4.3.

4.1 Approximate EWMA Stability Boundaries

Even for systems with relatively short metrology delays, finding analytical solutions to the stability boundary can be very difficult. For example, the stability boundary of a system with a delay of three runs requires the solution to a third order polynomial. To avoid this laborious task, a polynomial root solver was used

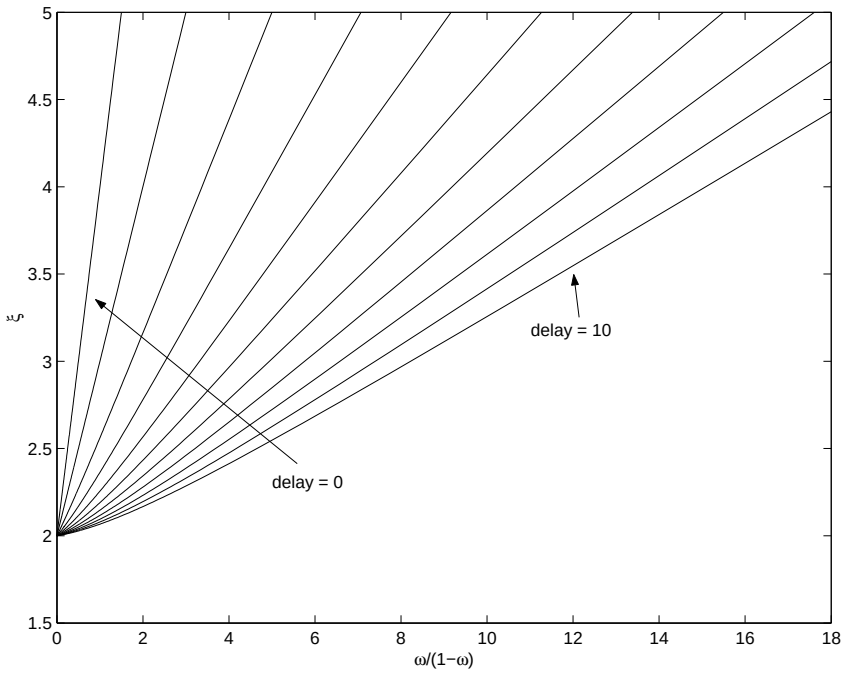

Figure 4.1: A nearly linear relationship between $\omega/(1-\omega)$ and ξ is observed

to find the stability boundaries of systems with metrology delays of longer than two runs as shown in Figure 2.4.

4.1.1 An Approximate SISO EWMA Stability Boundary

After using a polynomial solver to determine the stability regions for a limited number of metrology lags, it was observed that a nearly linear relationship exists between ξ and $\omega/(1-\omega)$ (i.e., the feasible region of ω is mapped to the RHP). Figure 4.1 shows a plot of ξ versus $\omega/(1-\omega)$ for the stability boundaries of the EWMA controller with metrology delays of zero through ten runs.

Figure 4.1 illustrates that the stability boundaries increase nearly linearly

with $1/(1-\omega)$. To understand the origins of this relationship between ξ and $\omega/(1-\omega)$, we will look closer at the stability region of a process with a metrology delay of two runs. First, the feasible region of ω is mapped to the RHP using the transformation,

$$\alpha = \frac{\omega}{1-\omega} \tag{4.1}$$

so that the feasible region for α is $0 \leq \alpha \leq \infty$. By substituting (4.1) into the analytical stability boundary in (2.18) we can see the relationship between α and ξ,

$$\xi = \frac{\alpha + 2 + \sqrt{5\alpha^2 + 8\alpha + 4}}{2}. \tag{4.2}$$

Next, differentiating ξ with respect to α yields

$$\frac{d\xi}{d\alpha} = -\frac{1}{2} + \frac{1}{4}\frac{10\alpha + 8}{\sqrt{5\alpha^2 + 8\alpha + 4}} \tag{4.3}$$

so we immediately see that the relationship between ξ and α is in fact not linear, as it may first appear. A plot of $\frac{d\xi}{d\alpha}$ versus α is shown in Figure 4.2 for processes with delays of two, three and four runs for the range of $\alpha = 0$ ($\omega = 0$) to $\alpha = 10$ ($\omega = 0.91$). Here we see that the derivative of ξ with respect to α converges to a constant very quickly. This explains why a plot of ξ versus α shows a nearly linear relationship.

To estimate the stability boundaries, the derivative of ξ with respect to α is broken up into two piecewise discontinuous parts: when ω is less than or equal to $\frac{1}{2}$ (i.e., when $\alpha \leq 1$) and when ω is greater than $\frac{1}{2}$ (i.e., when $\alpha > 1$). When $\omega \leq \frac{1}{2}$, $\frac{d\xi}{d\alpha}$ is approximated with a straight line and when $\omega > \frac{1}{2}$, $\frac{d\xi}{d\alpha}$ is approximated with the convergent value of $\frac{d\xi}{d\alpha}$ as ω approaches unity (i.e., $\alpha \to \infty$). In other words, ξ is

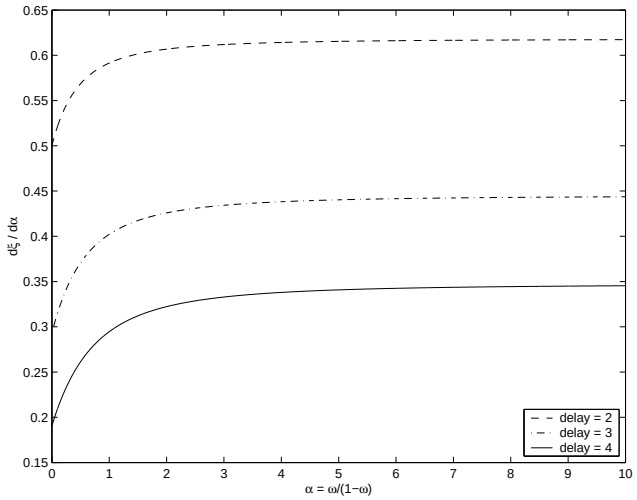

Figure 4.2: As α approaches infinity (i.e., as ω approaches unity), $\frac{d\xi}{d\alpha}$ converges to a constant

fit with a quadratic function for $0 \leq \alpha \leq \frac{1}{2}$ and a linear function for $\frac{1}{2} < \alpha$. These approximations are summarized in Table 4.1.

The approximate stability regions for delays of zero through four runs are plotted along with their respective analytical solutions in Figure 4.3. Here we see that the approximate solutions shown in Table 4.1 are, as a whole, very accurate. In the region, $0.6 \leq \omega \leq 0.9$, however, a small error develops between the analytical and approximate solutions. This is because small values of α are not well approximated by a constant slope (see Figure 4.2). It follows that as ω approaches unity, we see that the approximation becomes more accurate as the slope of ξ versus α converges to a constant. It is also worth noting that the ξ intercepts for $\alpha \leq 1$ in Table 4.1 are all exactly 2. This result is consistent with the results in Section 2.1.2.3 where it was shown that a sufficient condition for stability is that ξ must be between 0 and

Table 4.1: Approximate upper stability boundaries for delays of up to 10 runs, the lower stability boundaries are $\xi = 0$ for all metrology delays

	Stability Boundary	
delay	$0 \leq \alpha \leq 1$	$1 < \alpha$
0	$\xi = 2\alpha + 2$	$\xi = 2\alpha + 2$
1	$\alpha + 2$	$\alpha + 2$
2	$0.0471\alpha^2 + 0.517\alpha + 2$	$0.618\alpha + 1.944$
3	$0.0572\alpha^2 + 0.308\alpha + 2$	$0.445\alpha + 1.918$
4	$0.0545\alpha^2 + 0.203\alpha + 2$	$0.347\alpha + 1.908$
5	$0.0486\alpha^2 + 0.143\alpha + 2$	$0.284\alpha + 1.906$
6	$0.0423\alpha^2 + 0.105\alpha + 2$	$0.241\alpha + 1.906$
7	$0.0367\alpha^2 + 0.0808\alpha + 2$	$0.206\alpha + 1.910$
8	$0.0319\alpha^2 + 0.0639\alpha + 2$	$0.183\alpha + 1.912$
9	$0.0278\alpha^2 + 0.0517\alpha + 2$	$0.165\alpha + 1.914$
10	$0.0244\alpha^2 + 0.0427\alpha + 2$	$0.149\alpha + 1.917$

2, irrespective of the metrology delay.

The approximations in Table 4.1 are neither necessary nor sufficient conditions for stability, but rather numerical approximations for the necessary and sufficient stability boundaries. We can, however, derive a necessary condition by drawing a line starting at $\xi = 2$ on the ξ-axis and with the asymptotically convergent slopes listed in the third column of Table 4.1. If the model-mismatch is above this line, then the closed-loop system is guaranteed unstable. Likewise, a sufficient condition for stability can be derived by extending the asymptotically convergent line back to the ξ-axis. If the model-mismatch is below this line and greater than zero then the closed-loop system is guaranteed stable. This idea is illustrated in Figure 4.4 for a delay of two runs.

When the metrology delay is less than two runs, then ξ has a linear rela-

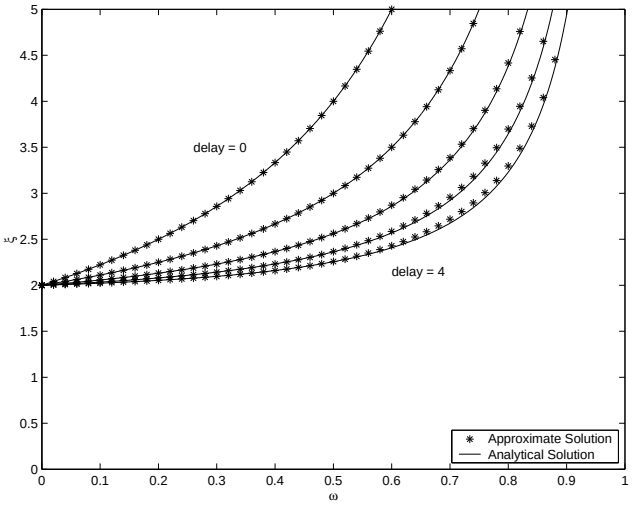

Figure 4.3: Stability region approximations of the EWMA controller for delays of zero through three runs

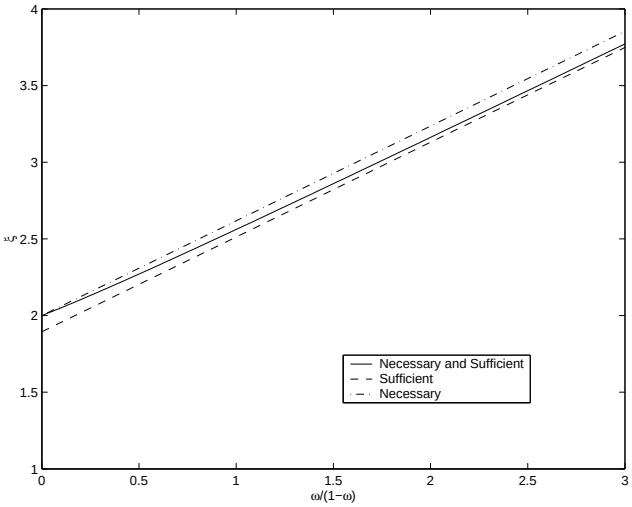

Figure 4.4: A necessary and a sufficient condition for stability for a delay of two runs

tionship with $\omega/(1-\omega)$ and the approximations in Table 4.1 are equivalent to the analytical necessary and sufficient conditions for stability. For a delay of two runs, the asymptotically convergent slope can be determined by taking the limit of (4.3) as α approaches infinity,

$$\lim_{\alpha\to\infty}\left\{-\frac{1}{2}+\frac{1}{4}\frac{10\alpha+8}{\sqrt{5\alpha^2+8\alpha+4}}\right\}=\frac{\sqrt{5}-1}{2}. \tag{4.4}$$

Next, using the basic linear form, $\xi = a_1\alpha + a_0$ and substituting (4.2) for ξ and the asymptotically convergent slope in (4.4) as a_1, we have

$$\frac{\alpha+2+\sqrt{5\alpha^2+8\alpha+4}}{2}=\frac{\sqrt{5}-1}{2}\alpha+a_0. \tag{4.5}$$

Solving for a_0 and taking the limit as α approaches infinity yields the ξ-intercept,

$$\lim_{\alpha\to\infty} a_0=\lim_{\alpha\to\infty}\left\{1+\frac{1}{2}\sqrt{5\alpha^2+8\alpha+4}-\frac{1}{2}\alpha\sqrt{5}\right\}=\frac{2\sqrt{5}+5}{5}. \tag{4.6}$$

Therefore, a necessary condition for stability is

$$0<\xi<\frac{\sqrt{5}-1}{2}\frac{\omega}{1-\omega}+2 \tag{4.7}$$

and a sufficient condition for stability is

$$0<\xi<\frac{\sqrt{5}-1}{2}\frac{\omega}{1-\omega}+\frac{2\sqrt{5}+5}{5}. \tag{4.8}$$

For delays of longer than two runs, the ξ-intercepts are determined numerically. The sufficient and the necessary (not to be confused with necessary *and* sufficient) conditions for stability are shown in Table 4.2 for delays of up to ten runs and these conditions are plotted alongside the analytical solutions for the stability boundaries in Figure 4.5.

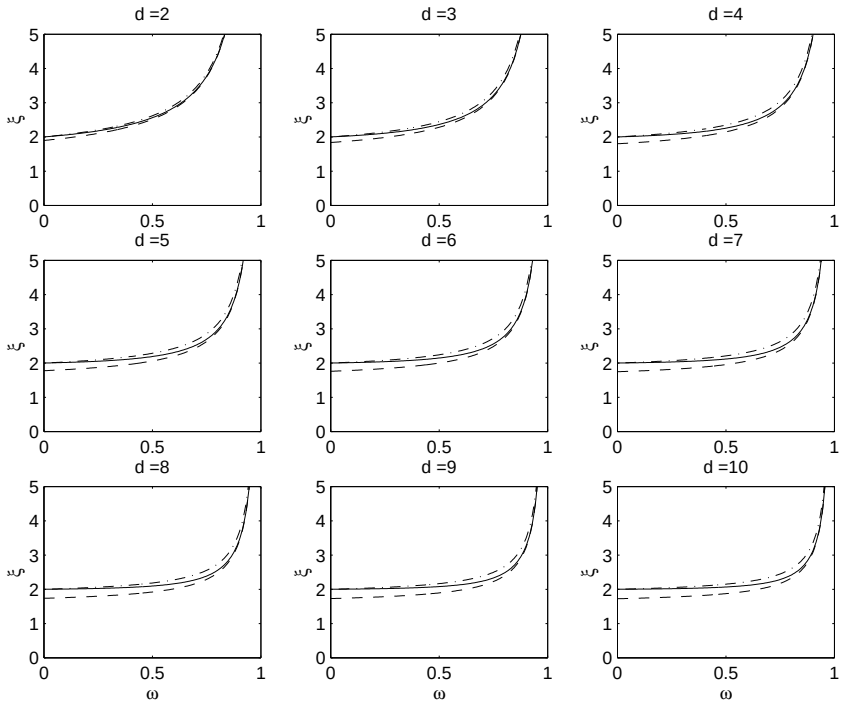

Figure 4.5: A necessary and a sufficient conditions for stability are plotted alongside the necessary and sufficient stability boundaries for delays of two through ten runs. The necessary condition and the sufficient condition is plotted with a dashed-dot and a dashed line, respectively.

Table 4.2: A necessary stability boundary and a sufficient stability boundary for delays of up to 10 runs

delay	Necessary	Sufficient
0	$\xi = 2\alpha + 2$	$\xi = 2\alpha + 2$
1	$\alpha + 2$	$\alpha + 2$
2	$0.618\alpha + 2$	$0.618\alpha + 1.89$
3	$0.445\alpha + 2$	$0.445\alpha + 1.83$
4	$0.347\alpha + 2$	$0.347\alpha + 1.80$
5	$0.284\alpha + 2$	$0.284\alpha + 1.77$
6	$0.241\alpha + 2$	$0.241\alpha + 1.75$
7	$0.206\alpha + 2$	$0.206\alpha + 1.74$
8	$0.183\alpha + 2$	$0.183\alpha + 1.73$
9	$0.165\alpha + 2$	$0.165\alpha + 1.72$
10	$0.149\alpha + 2$	$0.149\alpha + 1.72$

4.1.2 An Approximate MIMO EWMA Stability Boundary

In analogy to the SISO stability region in Section 4.1.1, an accurate numerical approximation can be derived for the MIMO EWMA stability boundaries. According to Theorem 2.2.3 the necessary and sufficient condition for closed loop stability is that all of the eigenvalues of the model mismatch matrix fall inside a complex stability boundary. We will derive an approximation of this boundary by first writing the eigenvalues of the model mismatch, λ, in complex polar form, $\lambda = re^{i\theta}$, where r is the complex modulus and θ is the phase angle [32]. By plotting r versus $\omega/(1-\omega)$ for a given phase angle we again observe a nearly linear relationship. This is shown in Figure 4.6 for delays of zero through ten runs and for phase angles of zero through $\frac{3\pi}{8}$.

Next, the slopes and intercepts of the lines in Figure 4.6 are plotted versus phase angle. The slopes are calculated using least squares regression for the range

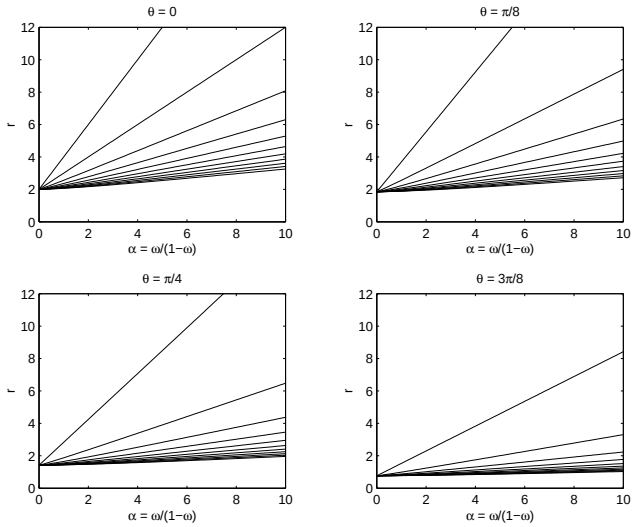

Figure 4.6: A nearly linear relationship between $\omega/(1-\omega)$ and complex modulus is observed for varying phase angles

$0 \leq \alpha \leq 10$ and are plotted in Figure 4.7. For the sake of simplicity, we have abandoned the piece-wise fitting used in Section 4.1.1. We know from Theorem 2.2.3 that the intercepts of the lines in Figure 4.6 follow the unit circle centered at {1,0} so that the intercepts are simply $2\cos\theta$ for any length of metrology delay. This is shown in Figure 4.8.

Figure 4.7 shows that the slope of the relationship between r and $\omega/(1-\omega)$ decreases according to a approximately linear relationship with the phase angle where the slope is zero when $\theta = \frac{\pi}{2}$ and similar to the SISO coefficient on $\omega/(1-\omega)$ shown Table 4.1 when the phase angle is zero. These results have an intuitive appeal for two reasons. First, when the phase angle is zero, the stability boundary of the MIMO EWMA controller should be equivalent to the SISO stability boundaries

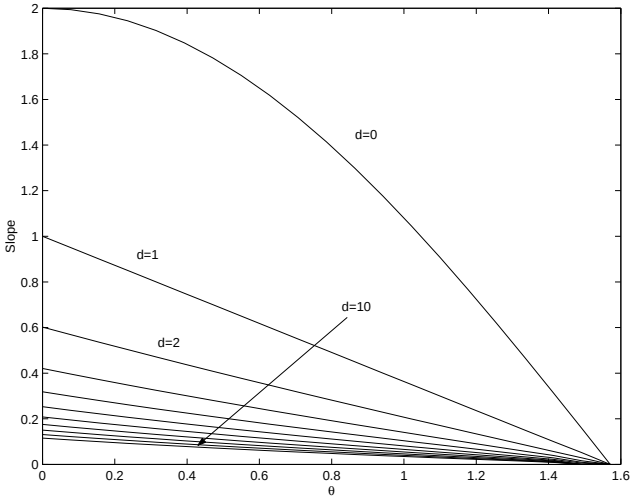

Figure 4.7: A nearly linear relationship is observed between the slope of r versus $\omega/(1-\omega)$ for varying phase angles

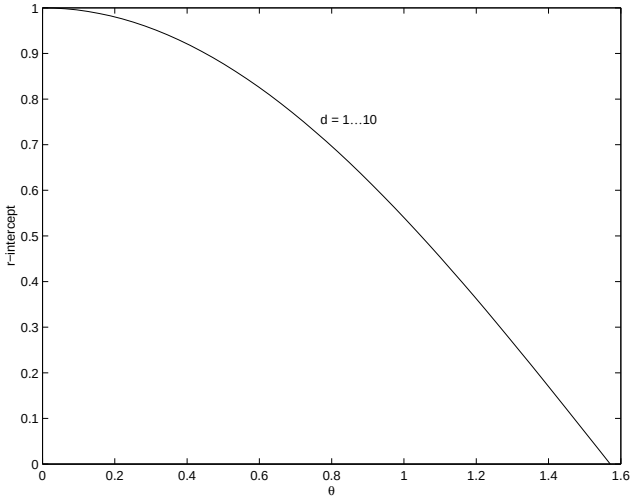

Figure 4.8: The intercepts of r versus α tracks the circle centered at {1,0} with unit radius for all metrology delays

shown in Table 4.1 with the respect to the eigenvalues of the model mismatch matrix, λ. The results are very close, with the differences resulting mainly from the use of the piecewise linear model in Table 4.1 while a simple linear model is used for the MIMO case. Next, we see that when the phase angle approaches $\frac{\pi}{2}$, the slope of λ versus α approaches zero. This too is intuitive in that the stability should be independent of ω when the eigenvalues of the model mismatch matrix are all completely imaginary. Finally, we see in Figure 4.7 that, for the case with no metrology delay, the slope of r with respect to θ tracks the circle centered at {1,0} on the complex plane with unit radius. This result is consistent with the analytical expression for the EWMA stability boundary in (2.29).

By combining the intercept and slope terms, an approximation for the complex stability boundary can be formed. These are show in Table 4.3 for metrology lags of zero through ten runs. The accuracy of the numerical approximation of the complex stability region is then demonstrated in Figure 4.6 where the numerical approximations and analytical solutions are plotted for systems with delays of one through four runs.

We see in Figure 4.9 that the approximation very closely matches the analytical solutions for a wide range of complex eigenvalues of the model mismatch matrix. This is truly a remarkable result considering the complexity and non-linearity of the MIMO stability boundaries. The structure of the numerical approximation allows us to accurately describe the stability boundaries using only one fitting parameter for each metrology delay. It could also be argued that the fitting parameters themselves could be approximated so that the stability boundaries could be described

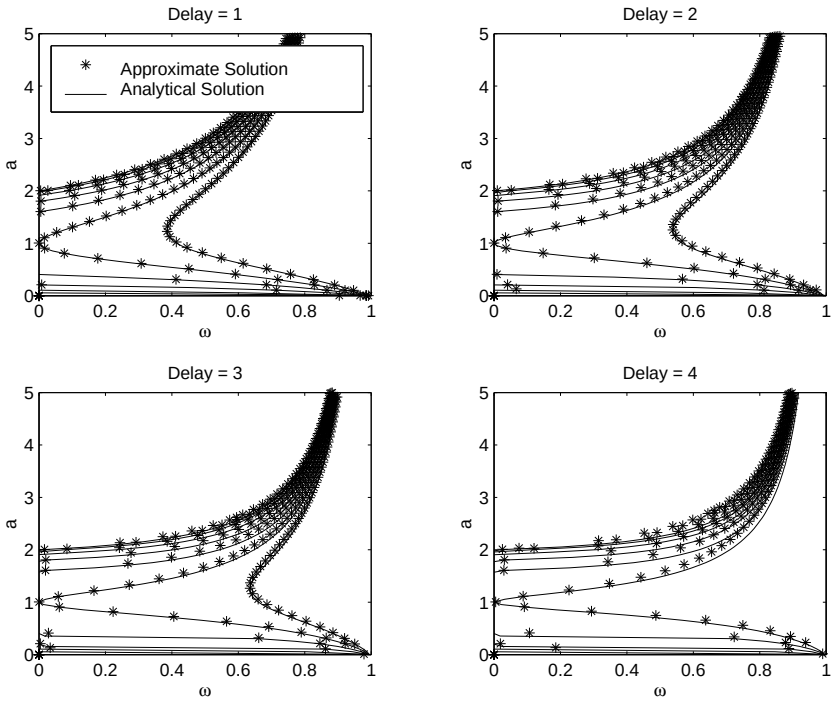

Figure 4.9: Approximate solution of the complex stability region for a process with metrology delays of one through four runs

using only two or three fitting parameters for any length of metrology delay.

4.2 Approximate Double-EWMA Stability Boundaries

A similar approach can be used to approximate the stability bounds of the double-EWMA controller. There are, however, two characteristics of the double-EWMA stability boundaries that should be addressed. First, unlike the EWMA stability boundaries, the double-EWMA controller can be unstable when the model mismatch is between zero and one. Therefore, a numerical approximation must be generated for both the upper and lower stability boundaries. Second, we see in Figure 3.3, for example, that the stability boundaries are discontinuous. It follows that a numerical approximation should account for these discontinuous shifts.

Table 4.3: Approximate complex stability regions of the EWMA controller for delays of up to 10 runs

Delay	Stability Region
0	$0 < r < 2\cos\theta\left(1+\frac{\omega}{1-\omega}\right)$
1	$0 < r < 2\cos\theta+\frac{\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
2	$0 < r < 2\cos\theta+\frac{0.602\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
3	$0 < r < 2\cos\theta+\frac{0.421\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
4	$0 < r < 2\cos\theta+\frac{0.383\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
5	$0 < r < 2\cos\theta+\frac{0.252\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
6	$0 < r < 2\cos\theta+\frac{0.208\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
7	$0 < r < 2\cos\theta+\frac{0.175\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
8	$0 < r < 2\cos\theta+\frac{0.150\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
9	$0 < r < 2\cos\theta+\frac{0.131\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$
10	$0 < r < 2\cos\theta+\frac{0.115\omega\left(1-\frac{2\theta}{\pi}\right)}{1-\omega}$

4.2.1 An Approximate SISO Double-EWMA Stability Boundary

Figure 4.10 shows that the double-EWMA stability boundaries, like the EWMA boundaries, can be well approximated by a piecewise linear relationship between ξ and α. As the length of the metrology delay increases, the number of active stability boundaries for the double-EWMA increases. This is clearly shown in Figure 4.10 where there are sharp, discontinuous shifts in the relationship between α and ξ. As the length of the metrology delay increases, however, many of the active boundaries show a gradual change in the α domain. It follows that, although the true stability boundaries may include multiple active regions, the boundaries can be well approximated using only two or three piecewise linear regions. We have chosen three active regions as a satisfactory approximation of the upper stability boundary and two piecewise regions for the lower boundary. These approximations are summarized in Tables 4.4 and 4.5, for the upper and lower stability boundaries, respectively.

The approximate stability boundaries for delays of zero through four runs is plotted in Figure 4.11.

4.2.2 An Approximate MIMO Double-EWMA Stability Boundary

In this section an approximate solution to the MIMO double-EWMA stability boundaries is derived. The approach must be different from the approximate complex EWMA stability boundary in Section 4.1.2 since the relationship between α and r is not well approximated with a straight line. An example of this is shown in Figure 4.12 for a delay of one run.

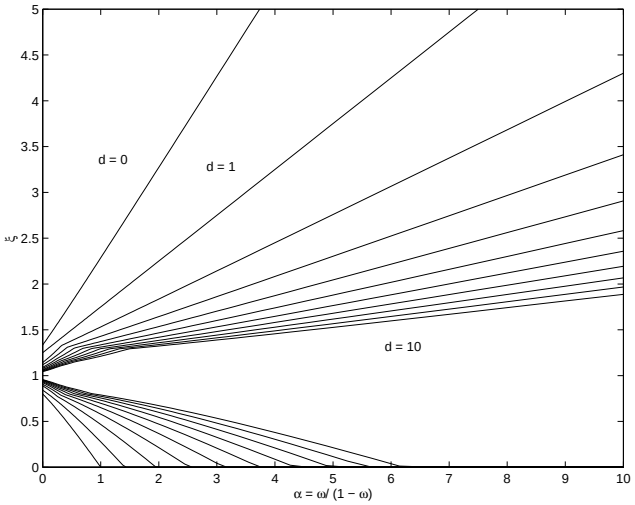

Figure 4.10: A nearly linear relationship between $\omega/(1-\omega)$ and ξ is observed for the upper and lower stability boundaries of the double-EWMA controller

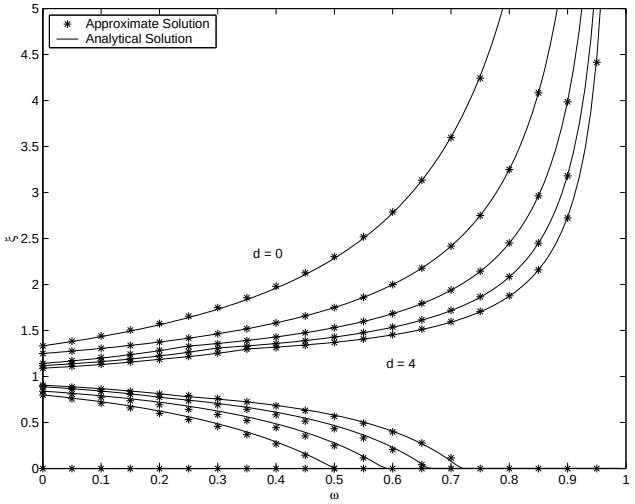

Figure 4.11: Stability region approximations of the double-EWMA controller for delays of zero through four runs

Table 4.4: Approximate lower stability boundaries of the double-EWMA controller for delays of up to 10 runs

delay	range	boundary	range	boundary	range	boundary
0	$0 \leq \alpha < \infty$	$\xi = 0$				
1	$0 \leq \alpha < 1$	$\xi = 0.8 - \alpha$	$1 \leq \alpha < \infty$	$\xi = 0$		
2	$0 \leq \alpha < 1.42$	$\xi = 0.841 - 0.594\alpha$	$1.42 \leq \alpha < \infty$	$\xi = 0$		
3	$0 \leq \alpha < 0.264$	$\xi = 0.889 - 0.459\alpha$	$0.264 \leq \alpha < 1.950$	$\xi = 0.888 - 0.456\alpha$	$1.950 \leq \alpha < \infty$	$\xi = 0$
4	$0 \leq \alpha < 0.318$	$\xi = 0.905 - 0.379\alpha$	$0.318 \leq \alpha < 2.534$	$\xi = 0.897 - 0.354\alpha$	$2.534 \leq \alpha < \infty$	$\xi = 0$
5	$0 \leq \alpha < 0.387$	$\xi = 0.923 - 0.336\alpha$	$0.387 \leq \alpha < 3.132$	$\xi = 0.905 - 0.289\alpha$	$3.132 \leq \alpha < \infty$	$\xi = 0$
6	$0 \leq \alpha < 0.466$	$\xi = 0.932 - 0.288\alpha$	$0.466 \leq \alpha < 3.739$	$\xi = 0.911 - 0.244\alpha$	$3.739 \leq \alpha < \infty$	$\xi = 0$
7	$0 \leq \alpha < 0.550$	$\xi = 0.941 - 0.255\alpha$	$0.550 \leq \alpha < 4.348$	$\xi = 0.917 - 0.211\alpha$	$4.348 \leq \alpha < \infty$	$\xi = 0$
8	$0 \leq \alpha < 0.642$	$\xi = 0.947 - 0.225\alpha$	$0.642 \leq \alpha < 4.952$	$\xi = 0.922 - 0.186\alpha$	$4.952 \leq \alpha < \infty$	$\xi = 0$
9	$0 \leq \alpha < 0.736$	$\xi = 0.952 - 0.201\alpha$	$0.736 \leq \alpha < 5.579$	$\xi = 0.926 - 0.166\alpha$	$5.579 \leq \alpha < \infty$	$\xi = 0$
10	$0 \leq \alpha < 0.832$	$\xi = 0.956 - 0.182\alpha$	$0.832 \leq \alpha < 6.194$	$\xi = 0.930 - 0.150\alpha$	$6.194 \leq \alpha < \infty$	$\xi = 0$

Table 4.5: Approximate upper stability boundaries of the double-EWMA controller for delays of up to 10 runs

delay	range	boundary	range	boundary
0	$0 \leq \alpha < \infty$	$\xi = 1.333 + \alpha$		
1	$0 \leq \alpha < \infty$	$\xi = 1.25 + 0.5\alpha$		
2	$0 \leq \alpha < 0.326$	$\xi = 1.143 + 0.570\alpha$	$0.326 \leq \alpha < \infty$	$\xi = 1.225 + 0.307\alpha$
3	$0 \leq \alpha < 0.425$	$\xi = 1.118 + 0.451\alpha$	$0.425 \leq \alpha < \infty$	$\xi = 1.210 + 0.219\alpha$
4	$0 \leq \alpha < 0.548$	$\xi = 1.091 + 0.384\alpha$	$0.548 \leq \alpha < \infty$	$\xi = 1.201 + 0.169\alpha$
5	$0 \leq \alpha < 0.689$	$\xi = 1.079 + 0.317\alpha$	$0.689 \leq \alpha < \infty$	$\xi = 1.195 + 0.138\alpha$
6	$0 \leq \alpha < 0.838$	$\xi = 1.067 + 0.272\alpha$	$0.838 \leq \alpha < \infty$	$\xi = 1.191 + 0.116\alpha$
7	$0 \leq \alpha < 0.992$	$\xi = 1.060 + 0.235\alpha$	$0.992 \leq \alpha < \infty$	$\xi = 1.187 + 0.0995\alpha$
8	$0 \leq \alpha < 1.155$	$\xi = 1.053 + 0.208\alpha$	$1.155 \leq \alpha < \infty$	$\xi = 1.185 + 0.0871\alpha$
9	$0 \leq \alpha < 1.320$	$\xi = 1.048 + 0.185\alpha$	$1.320 \leq \alpha < \infty$	$\xi = 1.183 + 0.0775\alpha$
10	$0 \leq \alpha < 1.481$	$\xi = 1.044 + 0.167\alpha$	$1.481 \leq \alpha < \infty$	$\xi = 1.182 + 0.0696\alpha$

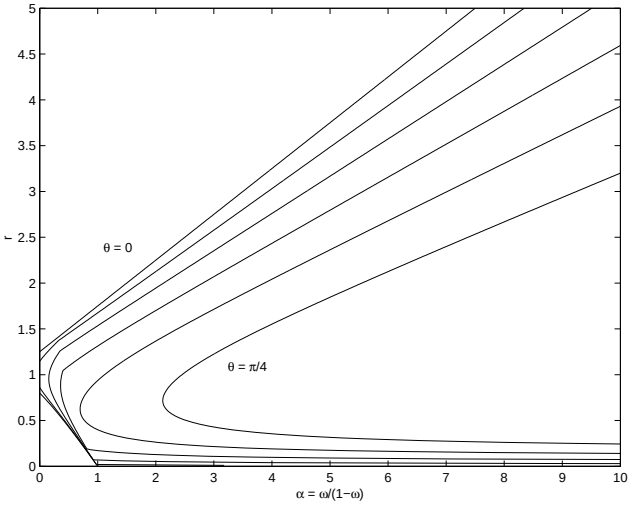

Figure 4.12: The complex stability boundary is not well approximated as a linear relationship with α. The relationship for a delay of one run is shown.

To approximate the double-EWMA complex stability boundaries, a method similar to that used to find a sufficient condition for stability in Section 3.2.2.3 is used. In this case, a sufficient condition for MIMO stability is determined by finding the largest circle that will fit inside the complex stability region. This circle is calculated by minimizing the maximum radius from a point inside the stability region to the complex stability boundary,

$$\min_{\lambda_0,\theta} \max_{r_0} r_0 = f(\lambda_0, \theta). \tag{4.9}$$

The minimax radius, r_0, is the Euclidean distance from a point, λ_0, which lies inside the complex stability region, to the complex stability boundary in (3.19),

$$r_0 = \left\| \lambda_0 - \left\{ 1 + \frac{z^d (z-\omega)^2}{(1-\omega)(d(1-\omega)+2)z + (1-\omega)^2(1+d)} \right\} \right\|_2 \tag{4.10}$$

and where the unit circle is written in polar notation, $z = e^{i\theta}$. We will refer to this sufficient condition for stability as the *sufficient stability disk*.

This problem is similar to the problem posed independently by Chrystal and Sylvester well over 100 years ago concerning finding the minimum radius which will encircle a group of points on a plane [21, 62]. The problem has received a considerable amount of attention and several algorithms exist which efficiently compute this minimax radius [26, 28, 49]. The details of these algorithms are outside the scope of this disseration but a survey of minimax location finding algorithms can be found in [35].

Let us consider a process with a delay of two runs and a tuning parameter of $\omega = 0.5$. Solving (4.9) using the algorithm of Elzinga and Hearn in [26] shows that the largest circle that can be inscribed inside the stability region is centered at $\lambda_0 = \{0.934, 0\}$ in the complex plane and has a radius of $r_0 = 0.506$. This is shown in Figure 4.13.

Figure 4.13 shows that the approximate MIMO stability boundaries are conservative as values close to the SISO stability boundaries (i.e., $\Im(\lambda) = 0$) fall well outside of the sufficient stability disk. However, the majority of the stable region is enclosed by this sufficient stability condition and this methodology will be used to find an approximate expression for the MIMO stability boundaries. We also see in Figure 4.13 that the stability boundary is symmetrical across the $\Re(\lambda)$ axis so that the center of the sufficient stability disk, λ_0, will have an imaginary component equal to zero.

Like the EWMA and SISO double-EWMA stability boundaries, both r_0 and

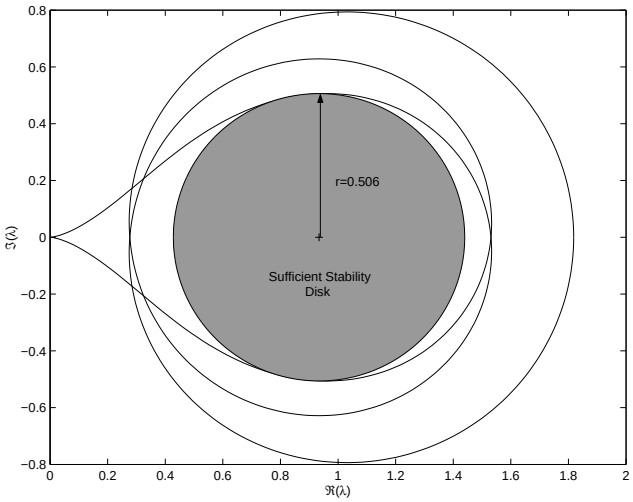

Figure 4.13: The complex boundary for $\omega = 0.5$ and $d = 2$. A sufficient condition for MIMO stability can be established by finding the largest circle that fits inside the complex stability boundary.

λ_0 follow an approximately piecewise piecewise linear relationship with α. This is plotted in Figures 4.14 and 4.15 for delays of zero through ten runs. The piecewise linear expressions approximating r_0 and λ_0 as a function of α are summarized in Tables 4.6 and 4.7, respectively.

Figure 4.16 shows a plot of the approximate complex double-EWMA stability boundaries along with the analytical solution for a system with a delay of one run. The approximation shows a satisfactory correlation with the analytical solution when $\lambda \approx 1$ and then shows poor correlation with increasing λ. This can be understood by looking at Figure 4.13. Here we see that the sufficient condition for stability is bounded by the imaginary component of the complex stability boundary and, as mentioned previously, there is a sizeable area in the stable region that is

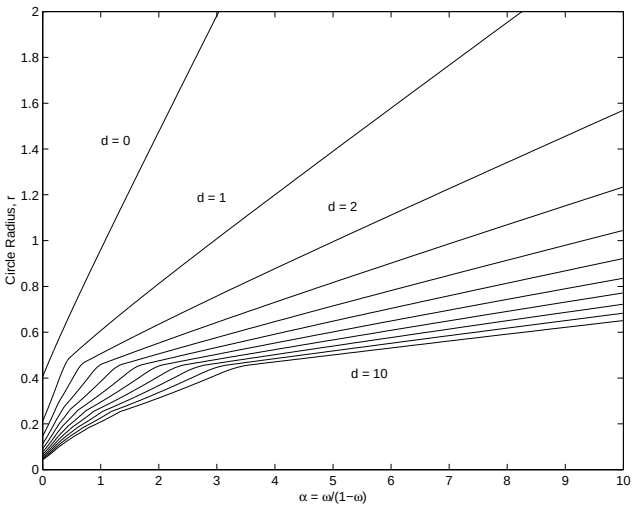

Figure 4.14: The radius of the largest circle that can be inscribed in the complex double-EWMA stability region follows an approximately piecewise linear relationship with α. Delays of 0 through 10 runs are shown.

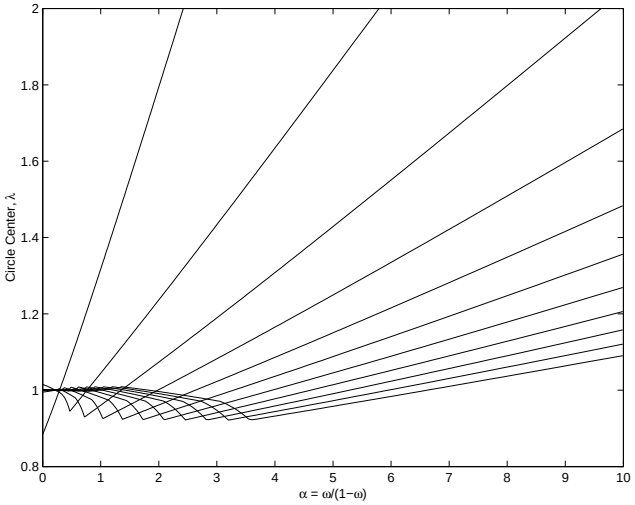

Figure 4.15: The center point of the circle of maximum radius in the complex double-EWMA stability region. Delays of 0 through 10 runs are shown.

Table 4.6: Piecewise linear approximation of the radius of the largest circle that can be inscribed in the complex double-EWMA stability region for delays of up to 10 runs

delay	range	boundary	range	boundary	range	boundary
0	$0 \leq \alpha < \infty$	$r_0 = 0.510\alpha + 0.438$				
1	$0 \leq \alpha < 0.47$	$r_0 = 0.652\alpha + 0.209$	$0.47 \leq \alpha < \infty$	$r_0 = 0.194\alpha + 0.415$		
2	$0 \leq \alpha < 0.28$	$r_0 = 0.534\alpha + 0.144$	$0.28 \leq \alpha < 0.73$	$r_0 = 0.424\alpha + 0.178$	$0.73 \leq \alpha < \infty$	$r_0 = 0.119\alpha + 0.393$
3	$0 \leq \alpha < 0.37$	$r_0 = 0.446\alpha + 0.111$	$0.37 \leq \alpha < 1.04$	$r_0 = 0.309\alpha + 0.158$	$1.04 \leq \alpha < \infty$	$r_0 = 0.0867\alpha + 0.379$
4	$0 \leq \alpha < 0.48$	$r_0 = 0.368\alpha + 0.0914$	$0.48 \leq \alpha < 1.38$	$r_0 = 0.239\alpha + 0.146$	$1.38 \leq \alpha < \infty$	$r_0 = 0.0680\alpha + 0.371$
5	$0 \leq \alpha < 0.62$	$r_0 = 0.309\alpha + 0.0788$	$0.62 \leq \alpha < 1.74$	$r_0 = 0.191\alpha + 0.141$	$1.74 \leq \alpha < \infty$	$r_0 = 0.0563\alpha + 0.365$
6	$0 \leq \alpha < 0.76$	$r_0 = 0.265\alpha + 0.0698$	$0.76 \leq \alpha < 2.11$	$r_0 = 0.159\alpha + 0.136$	$2.11 \leq \alpha < \infty$	$r_0 = 0.0479\alpha + 0.360$
7	$0 \leq \alpha < 0.90$	$r_0 = 0.230\alpha + 0.0631$	$0.90 \leq \alpha < 2.48$	$r_0 = 0.138\alpha + 0.130$	$2.48 \leq \alpha < \infty$	$r_0 = 0.0418\alpha + 0.357$
8	$0 \leq \alpha < 1.06$	$r_0 = 0.201\alpha + 0.0584$	$1.06 \leq \alpha < 2.85$	$r_0 = 0.122\alpha + 0.124$	$2.85 \leq \alpha < \infty$	$r_0 = 0.0370\alpha + 0.354$
9	$0 \leq \alpha < 1.21$	$r_0 = 0.179\alpha + 0.0542$	$1.21 \leq \alpha < 3.23$	$r_0 = 0.108\alpha + 0.122$	$3.23 \leq \alpha < \infty$	$r_0 = 0.0332\alpha + 0.352$
10	$0 \leq \alpha < 1.36$	$r_0 = 0.164\alpha + 0.0501$	$1.36 \leq \alpha < 3.61$	$r_0 = 0.0977\alpha + 0.119$	$3.61 \leq \alpha < \infty$	$r_0 = 0.0301\alpha + 0.350$

Table 4.7: Piecewise linear approximation of the center of the largest circle that can be inscribed in the complex double-EWMA stability region for delays of up to 10 runs

delay	range	boundary	range	boundary	range	boundary
0	$0 \leq \alpha < \infty$	$\lambda_0 = 0.485\alpha + 0.848$				
1	$0 \leq \alpha < 0.47$	$\lambda_0 = -0.0728\alpha + 1.016$	$0.47 \leq \alpha < \infty$	$\lambda_0 = 0.201\alpha + 0.840$		
2	$0 \leq \alpha < 0.28$	$\lambda_0 = 1.00$	$0.28 \leq \alpha < 0.73$	$\lambda_0 = -0.0843\alpha + 1.028$	$0.73 \leq \alpha < \infty$	$\lambda_0 = 0.119\alpha + 0.837$
3	$0 \leq \alpha < 0.37$	$\lambda_0 = 1.00$	$0.37 \leq \alpha < 1.04$	$\lambda_0 = -0.0662\alpha + 1.032$	$1.04 \leq \alpha < \infty$	$\lambda_0 = 0.0844\alpha + 0.831$
4	$0 \leq \alpha < 0.48$	$\lambda_0 = 1.00$	$0.48 \leq \alpha < 1.38$	$\lambda_0 = -0.0522\alpha + 1.034$	$1.38 \leq \alpha < \infty$	$\lambda_0 = 0.0645\alpha + 0.831$
5	$0 \leq \alpha < 0.62$	$\lambda_0 = 1.00$	$0.62 \leq \alpha < 1.74$	$\lambda_0 = -0.0429\alpha + 1.036$	$1.74 \leq \alpha < \infty$	$\lambda_0 = 0.0521\alpha + 0.830$
6	$0 \leq \alpha < 0.76$	$\lambda_0 = 1.00$	$0.76 \leq \alpha < 2.11$	$\lambda_0 = -0.0358\alpha + 1.037$	$2.11 \leq \alpha < \infty$	$\lambda_0 = 0.0436\alpha + 0.829$
7	$0 \leq \alpha < 0.90$	$\lambda_0 = 1.00$	$0.90 \leq \alpha < 2.48$	$\lambda_0 = -0.0300\alpha + 1.037$	$2.48 \leq \alpha < \infty$	$\lambda_0 = 0.0375\alpha + 0.828$
8	$0 \leq \alpha < 1.06$	$\lambda_0 = 1.00$	$1.06 \leq \alpha < 2.85$	$\lambda_0 = -0.0275\alpha + 1.039$	$2.85 \leq \alpha < \infty$	$\lambda_0 = 0.0330\alpha + 0.827$
9	$0 \leq \alpha < 1.21$	$\lambda_0 = 1.00$	$1.21 \leq \alpha < 3.23$	$\lambda_0 = -0.0246\alpha + 1.040$	$3.23 \leq \alpha < \infty$	$\lambda_0 = 0.0292\alpha + 0.827$
10	$0 \leq \alpha < 1.36$	$\lambda_0 = 1.00$	$1.36 \leq \alpha < 3.61$	$\lambda_0 = -0.0223\alpha + 1.040$	$3.61 \leq \alpha < \infty$	$\lambda_0 = 0.0263\alpha + 0.827$

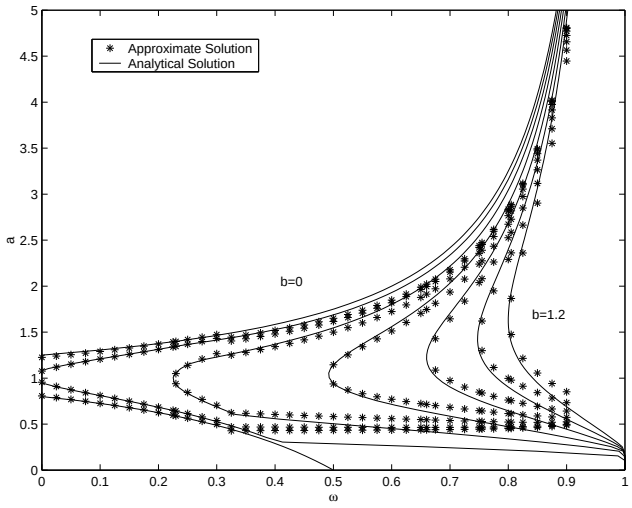

Figure 4.16: The approximate complex double-EWMA stability boundary for a delay of one run

not captured inside the sufficient stability disk near the real (i.e., SISO) stability boundary. It follows that when the real part of the eigenvalues of $\boldsymbol{\xi}$ are close to unity, then the instabilities must result from the complex components of the eigenvalues. Since the sufficient stability circle is close to (or touches) the imaginary component of the stability region, we see a closer approximation to the stability region when λ is close to unity. Then, because of the conservative nature of the approximation, we see a poor fit as λ deviates from unity. Nevertheless, the approximate MIMO double-EWMA stability boundaries correspond to a sufficient condition for closed-loop stability.

4.3 Conclusions

A fundamental concern in choosing the tuning parameter of the EWMA and double-EWMA run-to-run controllers is the closed loop stability of the process when there is uncertainty in the process model and feedback delay. It was shown in Chapters 2 and 3 that the solution to the stability boundaries of the EWMA and double-EWMA controllers are often prohibitively complicated, especially when the eigenvalues of the model mismatch matrix are complex and there are extended metrology delays. This chapter has reduced the complexity of the stability boundary problem by introducing a numerical approximation technique. Approximate stability boundaries have been derived for delays of up to ten runs for the SISO and MIMO EWMA and double-EWMA run-to-run controllers and the accuracy of the approximations are demonstrated by comparing the results with the analytical solutions.

Chapter 5

The Transient Performance of the EWMA Controller

In Chapter 2 we looked at the effects of metrology delay and model mismatch on the stability of the EWMA run-to-run controller. In addition to affecting the stability boundaries, the presence of metrology lag and model mismatch can result in a detrimental impact on the performance of a process under run-to-run control. In this chapter the closed-loop performance requirements for processes with delays of zero through two runs are derived using modified stability analysis tools. Several simulations are provided to illustrate the importance of considering metrology delay and model mismatch implicitly in controller performance and tuning.

The outline of the chapter is as follows. The conditions for meeting closed-loop transient performance requirements for the SISO EWMA controller are derived in Section 5.1. The MIMO EWMA transient performance conditions are then derived in Section 5.2. Some concluding remarks are then given in Section 5.3.

5.1 Performance of the SISO EWMA Controller

In this section the closed-loop performance criteria for the SISO EWMA controller are derived. The performance boundaries for the SISO EWMA controller without metrology delay is provided in Section 5.1.1 followed by the case with metrology delay in Section 5.1.2. A simulation illustrating these performance conditions

is provided in Section 5.1.3.

It was shown in (2.9) that the characteristic equation of the closed-loop SISO EWMA controller is

$$f(z) = z^{d+1} - z^d\omega - (1-\xi)(1-\omega) = 0, \tag{5.1}$$

where ξ was defined in (1.21) to be the ratio of the true process gain, β, and the estimated gain, b, $\xi = \frac{\beta}{b}$. It is well known that the closed-loop system is stable when all of the roots of the characteristic equation (5.1) are inside the unit circle. Furthermore, requirements on closed-loop *performance* can be established by solving for the conditions where the roots of (5.1) are constrained to a region smaller than the unit circle [2, 3]. This region is illustrated in Figure 5.1 and will be referred to as the *performance region*. In other words, the controller error is guaranteed to decay with the rate of ρ^t where t is the number of control steps if the roots of (5.1) all fall within the circle centered at {0,0} on the complex plane with a radius of ρ.

The procedure for determining the EWMA closed-loop performance boundaries will be demonstrated by considering systems with metrology lags of zero, one, and two runs.

5.1.1 SISO Performance Without Metrology Delay

For a process without metrology delay, (5.1) becomes

$$f(z) = z - \omega - (1-\xi)(1-\omega) = 0. \tag{5.2}$$

It is easy to see from (5.2) that the process will meet the performance criterion when

$$\frac{1-\rho}{1-\omega} \leq \xi \leq \frac{1+\rho}{1-\omega}. \tag{5.3}$$

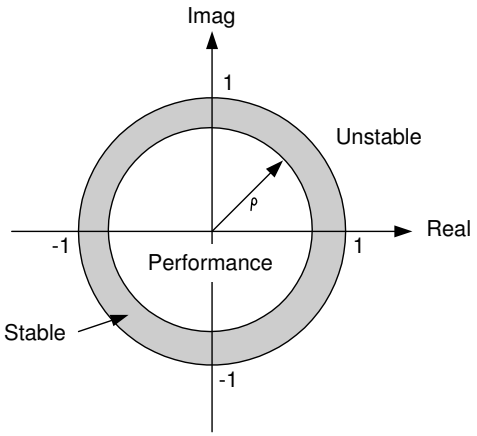

Figure 5.1: The performance region is defined as a circle with radius, $\rho < 1$, on the complex plane

The performance boundaries for $\rho = 0.2 \ldots 1.0$ are shown in Figure 5.2. When $\rho = 1$ the performance region is equivalent to the SISO stability region so that (5.3) is the same as the closed-loop stability boundary shown in (2.3) and shown in Figure 2.1. In addition, we see in Figure 5.2 that for a process with no model mismatch (i.e., $\xi = 1$), the highest performance is achieved when the EWMA tuning parameter factor, ω, is zero. Figure 5.2 also shows that for an underestimated model gain (i.e., $\xi \geq 1$), the controller tends to be too aggressive. It is, therefore, necessary to increase ω to improve the controller's performance (thus making the controller less aggressive). This is a reasonable result since with a large ξ, the controller will make large control moves based on an inaccurate process model. This results in poor performance. It follows that by increasing ω, the controller will make less aggressive control moves which leads to better performance. It could equivalently be stated that Figure 5.2 shows that for a small value of ω,

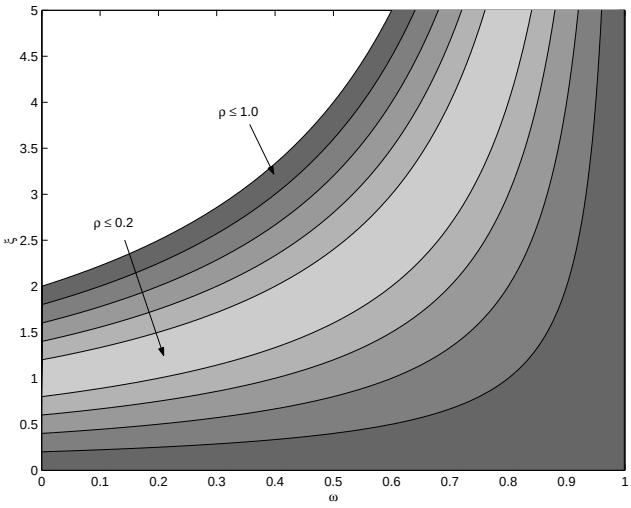

Figure 5.2: The performance region of the SISO EWMA system without metrology delay

it is necessary to decrease the plant-model mismatch in order to achieve a desired closed-loop performance. For a conservative choice of ω (i.e., large ω), ξ needs to be large to be more aggressive. Generally speaking, it is the value of $(1-\omega)\xi$ that determines the transient performance of the controller.

5.1.2 SISO Performance With Metrology Delay

5.1.2.1 A Delay of One Run

For a process with a metrology delay of one run, (5.1) becomes

$$f(z) = z^2 - z\omega - (1-\xi)(1-\omega) = 0. \tag{5.4}$$

To find the conditions that that guarantee that all of the roots of (5.4) meet the performance requirement, the we will follow the same procedure as the EWMA stability analysis in Section 2.1.2. First, the inside of the performance region is

mapped to the left half plane (LHP) using a bilinear transformation similar to (2.10),

$$z = \rho \left(\frac{1+w}{1-w} \right), \tag{5.5}$$

so that solving for the performance region boundaries is reduced to finding the conditions where roots for w are in the LHP. The characteristic equation of a process with a delay of one run becomes

$$f(w) = a_2 w^2 + a_1 w + a_0 = 0 \tag{5.6}$$

where

$$\begin{aligned} a_0 &= \rho^2 - \omega\rho - \xi\omega + \omega + \xi - 1 \\ a_1 &= 2(\rho^2 + \xi\omega - \xi - \omega + 1) \\ a_2 &= \rho^2 + \omega\rho - \xi\omega + \omega + \xi - 1. \end{aligned}$$

The roots of (5.6) are all in the LHP when the coefficients on w have the same sign. Solving this system of inequalities yields the performance criterion,

$$1 - \frac{\rho(\rho-\omega)}{1-\omega} \leq \xi \leq 1 + \frac{\rho^2}{1-\omega} \tag{5.7}$$

which is equivalent to the stability criterion in (2.12) when $\rho = 1$. The performance regions for $\rho = 0.2 \ldots 1.0$ are shown in Figure 5.3.

Figure 5.3 shows that, unlike the undelayed case shown in Figure 5.2, introducing metrology lag into the system places an upper limit on the closed-loop performance. More specifically, for the undelayed case, as long as process gain is underestimated (i.e., $\xi > 1$), then any performance requirement can be reached by choosing an appropriate ω. For the delayed case, on the other hand, achievable performance is limited by the plant-model mismatch. For example, when $\xi = 1.8$,

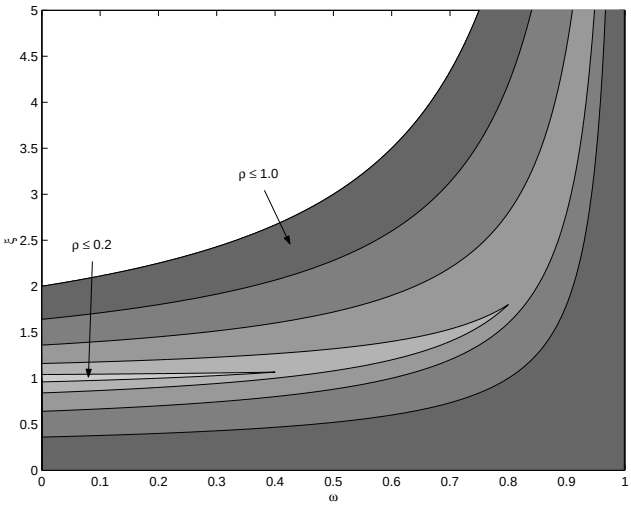

Figure 5.3: The performance region of the SISO EWMA system with a metrology delay of one run

the process will have at least one closed-loop pole greater than 0.4, irrespective of the choice of ω.

Rearranging (5.7) leads to

$$-\rho^2 \leq (1-\omega)(1-\xi) \leq \rho(\rho-\omega). \tag{5.8}$$

For an overestimated gain ($\xi < 1$), the controller is guaranteed stable for any ω. However, (5.8) mandates that $\rho \geq \omega$, which means that the closed-loop performance in this case is limited by the choice of ω. For a given ω and $\xi < 1$, (5.8) requires that

$$\rho \geq \frac{\omega + \sqrt{\omega^2(1-\xi)(1-\omega)}}{2} \tag{5.9}$$

which gives the exact lower bound for ρ. For an underestimated gain ($\xi > 1$), (5.8)

can be rewritten as

$$-\rho(\rho-\omega) \leq (\xi-1)(1-\omega) \leq \rho^2 \tag{5.10}$$

which gives a lower bound for ρ as

$$\min\left(\sqrt{(\xi-1)(1-\omega)}, \frac{\omega+\sqrt{\omega^2-(\xi-1)(1-\omega)}}{2}\right). \tag{5.11}$$

This is the performance limitation for a given ω and $\xi > 1$.

5.1.2.2 A Delay of Two Runs

By extending the delay to two runs, the characteristic equation in (5.1) becomes

$$f(z) = z^3 - z^2\omega - (1-\xi)(1-\omega) = 0$$

and after mapping the inside of the performance region to the LHP, the characteristic equation becomes

$$f(w) = a_3w^3 + a_2w^2 + a_1w + a_0 = 0 \tag{5.12}$$

where

$$\begin{aligned} a_0 &= \rho^3 - \omega\rho^2 + \xi - \xi\omega + \omega - 1 \\ a_1 &= 3\rho^3 - \omega\rho^2 - 3\omega - 3\xi + 3\xi\omega + 3 \\ a_2 &= 3\rho^3 + \omega\rho^2 + 3\omega + 3\xi - 3\xi\omega - 3 \\ a_3 &= \rho^3 + \omega\rho^2 - \omega - \xi + \xi\omega + 1. \end{aligned}$$

Applying the Routh-Hurwitz stability criterion on (5.12) yields

$$H = \begin{bmatrix} a_3 & a_1 \\ a_2 & a_0 \end{bmatrix}$$

and solving the system of inequalities which guarantees all of the principal minors of H are greater than zero results in the closed-loop performance region,

$$1 - \frac{\rho^2(\rho-\omega)}{1-\omega} \leq \xi \leq 1 - \frac{1}{2}\frac{\rho^2(\omega-\sqrt{\omega^2+4\rho^2})}{1-\omega}, \tag{5.13}$$

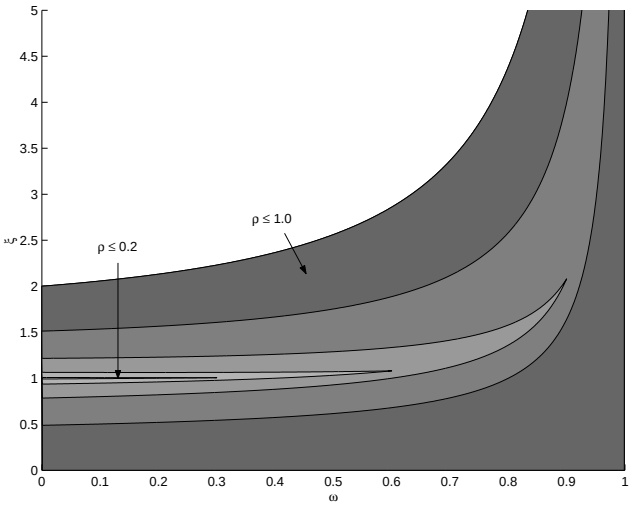

Figure 5.4: The performance region of the SISO EWMA system with a metrology delay of two runs

which is shown in Figure 5.4.

By comparing Figure 5.4 with Figures 5.2 and 5.3 we see that, not surprisingly, increasing the metrology delay further limits the performance of the controller. This same procedure can be used to find the performance boundaries for systems with larger metrology delays.

5.1.3 Simulation of a SISO Process

To demonstrate the importance of considering metrology delay and model mismatch implicitly in choosing the EWMA tuning parameters, we will look at a simple SISO simulation. The parameter constants for the simulation are as follows: $T = 0, \nu = 0, \beta = 4.0$ and has normally distributed noise with unit variance. Through a designed experiment or from process data, an approximation for the

Table 5.1: The maximum distance from the origin to the closed-loop poles, ρ

Delay	$\omega = 0.2$	$\omega = 0.5$	$\omega = 0.8$
0	0.600	0.00	0.600
1	0.894	0.707	0.447
2	0.961	0.872	0.701

model gain is determined to be $b = 2.0$ corresponding to a plant-model mismatch of $\xi = 2.0$. We will look at the step responses of processes with metrology delays of zero, one and two runs and with discounting factors of $\omega = 0.2$, 0.5, and 0.8. The maximum distances from the origin of the system poles for the simulations are summarized in Table 5.1.

Figure 5.5 shows the impact that metrology delay has on an EWMA controlled process. The first column in Figure 5.5 shows the step response of a process in the absence of metrology delay. In this case, $\omega = 0.2$ (top left) and $\omega = 0.8$ (bottom left) are on the $\rho = 0.6$ iso-performance line. Although the poles of these two systems are an equal distance from the origin, the system with $\omega = 0.2$ is underdamped so oscillations are observed. It could be argued that because of these oscillations, $\omega = 0.2$ and $\omega = 0.8$ do not have equivalent performance and that $\omega = 0.8$ has the best overall performance. Nevertheless, damping of the step disturbances of both simulations ($\omega = 0.2$ and $\omega = 0.8$) occur at the same rate so that the magnitude of the deviation from target during the transient period are equivalent for both simulations.

Simulations of a step response for systems with metrology delays of one and two runs are shown in the second and third columns of Figure 5.5, respectively.

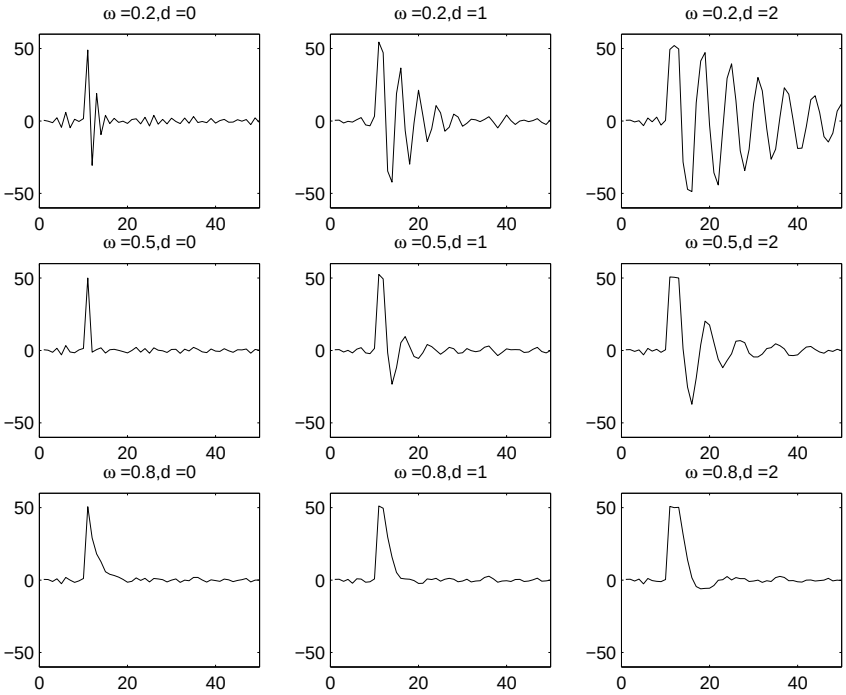

Figure 5.5: Simulations of a SISO EWMA controller with $\xi = 2$; $\omega = 0.2, 0.5, 0.8$; and delays of 0, 1, and 2 runs

Systems with metrology delay have an additional pole for each lag in the process metrology. These additional poles slow the response of the controller to disturbances so that the performance regions are more accurately described as the worst-case performance regions. That is, the performance regions shown in Figures 5.3 and 5.4 show the boundaries of the maximum system pole, but say nothing about the remaining poles. This is illustrated in Figure 5.5; although the process with $d = 1$ and $\omega = 0.5$ (the middle plot) has a larger maximum pole than the process with $d = 2$ and $\omega = 0.8$ (lower right), the two unaccounted for poles in the later process result in a slower response to the step disturbance. Nevertheless, the magnitude of the largest closed-loop pole effectively operates as a boundary of the worst-case transient response to a step disturbance.

5.2 Performance of the MIMO EWMA Controller

We showed in Chapter 2 that the necessary and sufficient condition for the stability of the MIMO EWMA controller is that all of the eigenvalues of the model mismatch matrix, $\boldsymbol{\xi}$, fall inside a complex stability region (the plant model mismatch matrices for the difference MIMO EWMA controller formulations are listed in equations (2.22) through (2.25)). It follows that the complex stability region is equivalent to the SISO stability region with respect to the eigenvalues of the model mismatch matrix. In other words, when the eigenvalues of $\boldsymbol{\xi}$ are all real, then the necessary and sufficient condition for MIMO stability is that all of the eigenvalues must fall inside the SISO stability region. This result can be extended to the performance regions of the MIMO system. That is, closed-loop performance is guaranteed

for the MIMO case when all of the eigenvalues of $\boldsymbol{\xi}$ fall inside a complex stability region. When all of the eigenvalues of the model mismatch matrix are real, then the performance requirement is that all of the these eigenvalues must fall inside the SISO performance regions in (5.3), (5.7), and (5.13) for systems with no delay, a delay of one run, and a delay of two runs, respectively.

Starting with the MIMO EWMA closed-loop realization derived in (2.30)

$$\hat{\nu}_{t+d+1} = \boldsymbol{\omega}\hat{\nu}_{t+d} + (I - \boldsymbol{\omega})(I - \boldsymbol{\xi})\hat{\nu}_t + (I - \boldsymbol{\omega})\nu + \varepsilon_{t+1} \tag{5.14}$$

where $\boldsymbol{\omega}$ is a diagonal matrix containing the EWMA filter constants for each of the m disturbance channels. Using the discounted RLS regression argument in Section 2.2.1, the diagonal elements of $\boldsymbol{\omega}$ can be chosen to be equivalent to one another. After this simplification, the z-transform of (2.30) becomes

$$z^{d+1} - z^d - (1 - \lambda_j)(1 - \omega) = 0 \qquad \forall j = 1, 2, \ldots, m. \tag{5.15}$$

To ensure a minimum closed-loop performance, we need the roots of (5.15) to fall within a radius, $\rho < 1$ for all $j = 1, 2, \ldots, m$. Similar to the SISO performance case in Section 5.1, the bilinear transform in (5.5) can be applied to (5.15) which gives rise to

$$\left(\rho\frac{1+w}{1-w}\right)^{d+1} - \omega\left(\rho\frac{1+w}{1-w}\right)^{d} - (1 - \lambda_j)(1 - \omega) = 0. \tag{5.16}$$

The characteristic equation in (5.16) will be used to find the performance boundaries for metrology delays of zero through two runs.

5.2.1 MIMO Performance Without Metrology Delay

Starting with (5.16) for a system without metrology delay, the generalized Routh-Hurwitz stability criterion detailed in Section 2.2.2 is used to find the closed-loop performance boundaries. Substituting $\lambda_j = \mathrm{a} + \mathrm{bi}$ and $w = wi$ yields the complex polynomial,

$$\begin{aligned} \Re\{f(wi)\} &= \mathrm{b}(1-\omega)w + \rho - 1 + \mathrm{a}(1-\omega) \\ \Im\{f(wi)\} &= (\rho + 1 - \mathrm{a}(1-\omega))w + \mathrm{b}(1-\omega) \end{aligned} \tag{5.17}$$

so that the ∇_{2n} matrix is

$$\nabla_{2n} = \begin{bmatrix} b_1 & b_0 \\ a_1 & a_0 \end{bmatrix}$$

where

$$\begin{aligned} b_0 &= \rho + 1 - \mathrm{a}(1-\omega) & a_0 &= \mathrm{b}(1-\omega) \\ b_1 &= \mathrm{b}(1-\omega) & a_1 &= \rho - 1 + \mathrm{a}(1-\omega). \end{aligned}$$

Solving the system of inequalities such that all of the principal minors are greater than zero yields the performance region,

$$\begin{aligned} \frac{|\lambda_j|^2 - \mathrm{a} - \sqrt{\rho^2|\lambda_j|^2 - \mathrm{b}^2}}{|\lambda_j|^2} &\leq \omega \\ &\leq \frac{|\lambda_j|^2 - \mathrm{a} + \sqrt{\rho^2|\lambda_j|^2 - \mathrm{b}^2}}{|\lambda_j|^2} \end{aligned} \tag{5.18}$$

which is equivalent to the result in [63] and (2.29) when $\rho = 1$.

The complex performance regions for the EWMA controller without metrology delay are shown in Figure 5.6. Here we see that the performance regions shrink and are pushed up and to the right as the complex part of λ_j increases. When the real part λ_j is close to unity then the imaginary part of λ_j becomes more dominant. It follows that the real part of λ_j must be increased to reach a satisfactory

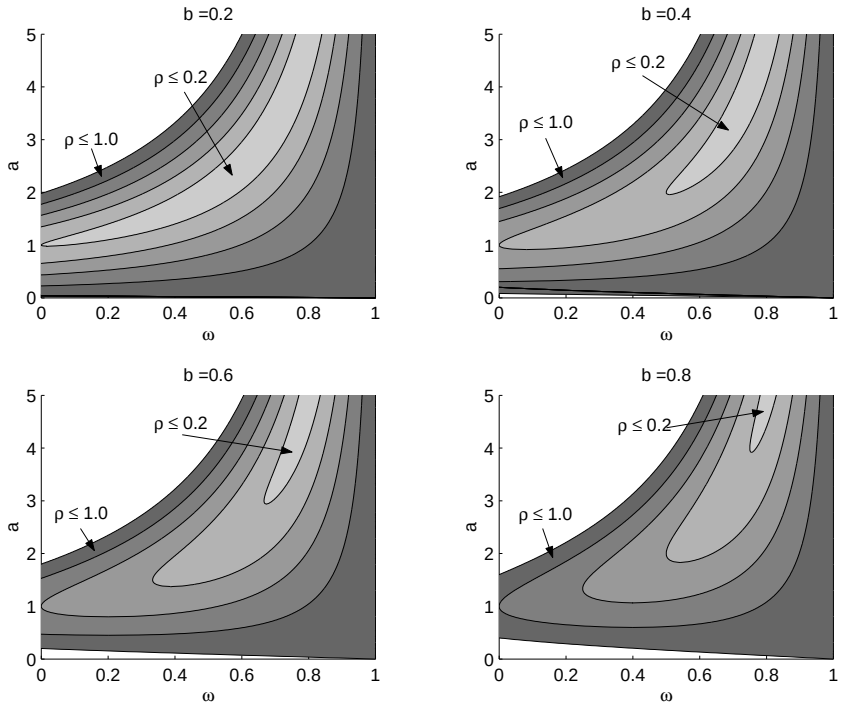

Figure 5.6: The complex performance regions for systems without metrology delay, b = 0.2 . . . 0.8

closed-loop performance. But, as with the SISO case, a larger λ_j means that a less aggressive forgetting factor is necessary to achieve a desired performance. Therefore, we observe the performance regions moving up and to the right with increasing complexity in λ_j. This is a truly surprising result as it implies that if the model mismatch has complex eigenvalues, it may be desirable to *increase* the real component of the mismatch matrix to achieve a desired closed-loop performance.

5.2.2 MIMO Performance With Metrology Delay

By following the same procedure as Section 5.2.1, the performance regions for the MIMO EWMA controller with metrology delay can be derived. We will look at the cases with metrology delays of one and two runs.

5.2.2.1 A Delay of One Run

Starting with the characteristic equation in (5.16) for $d = 1$ and substituting $w = iw$ and $\lambda = \mathrm{a} + \mathrm{b}i$ yields the complex polynomial,

$$\begin{aligned} \Re\{f(iw)\} &= a_2 w^2 + a_1 w + a_0 \\ \Im\{f(iw)\} &= b_2 w^2 + b_1 w + b_0 \end{aligned} \tag{5.19}$$

where

$$\begin{aligned} a_0 &= (1+\mathrm{a})(1-\omega) + \rho(\rho-\omega) \\ a_1 &= 2\mathrm{b}(1-\omega) \\ a_2 &= (1-\mathrm{a})(1-\omega) + \rho(\rho+\omega) \end{aligned}$$

$$\begin{aligned} b_0 &= \mathrm{b}(1-\omega) \\ b_1 &= 2\left(\rho^2 + (1-\mathrm{a})(1-\omega)\right) \\ b_2 &= -\mathrm{b}(1-\omega) \end{aligned}$$

and the ∇ matrix is

$$\nabla = \begin{bmatrix} b_2 & b_1 & b_0 & 0 \\ a_2 & a_1 & a_0 & 0 \\ 0 & b_2 & b_1 & b_0 \\ 0 & a_2 & a_1 & a_0 \end{bmatrix}.$$

Solving the system of inequalities such that the principal minors of ∇ are greater than zero[1] yields the performance regions shown in Figure 5.7.

Figure 5.7 shows that when eigenvalues of the plant model mismatch matrix are complex, the closed-loop performance degrades significantly. For example,

[1] These solutions were found numerically

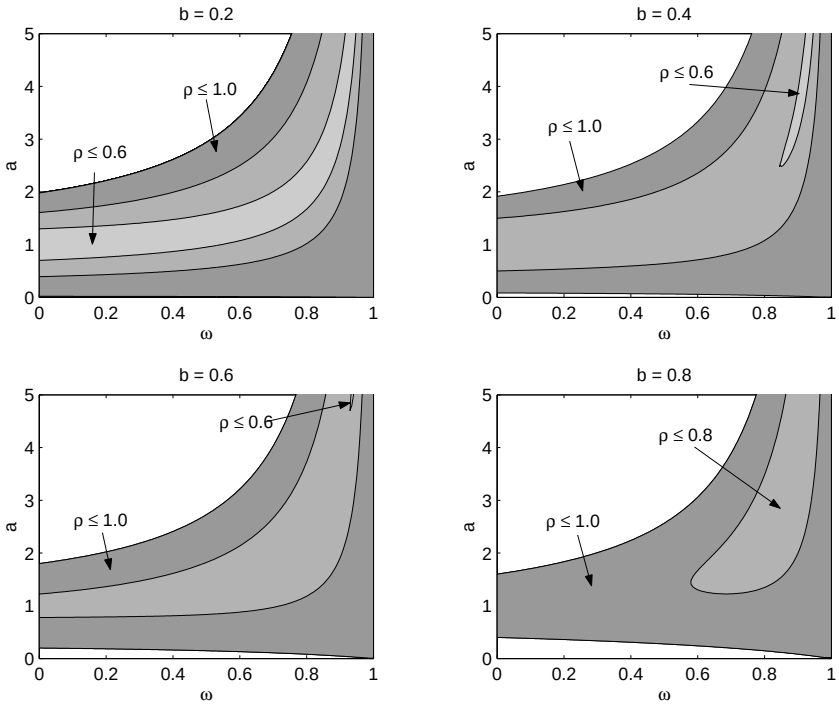

Figure 5.7: The complex performance regions for a system with metrology delay of one run, b = 0.2 . . . 0.8

Figure 5.7 shows that for a $= 0 \ldots 5$, and with imaginary part b ≥ 0.2, there are no values of ω in the feasible region $0 \leq \omega \leq 1$ that will guarantee that all of the closed-loop poles are inside a circle with a radius of $\rho = 0.4$.

5.2.2.2 A Delay of Two Runs

The complex characteristic equation for a process with a delay of two runs is

$$\begin{aligned} \Re\{f(iw)\} &= a_3 w^3 + a_2 w^2 + a_1 w + a_0 \\ \Im\{f(iw)\} &= b_3 w^3 + b_2 w^2 + b_1 w + b_0 \end{aligned} \tag{5.20}$$

where

$$\begin{gathered} a_0 = (\mathrm{a}-1)(1-\omega) + \rho^2(\rho - \omega) \\ a_1 = 3\mathrm{b}(1-\omega) \\ a_2 = -3(\mathrm{a}-1)(1-\omega) - \rho^2(3\rho + \omega) \\ a_3 = -\mathrm{b}(1-\omega) \end{gathered}$$

$$\begin{gathered} b_0 = \mathrm{b}(1-\omega) \\ b_1 = -3(\mathrm{a}-1)(1-\omega) + \rho^2(3\rho - \omega) \\ b_2 = -3\mathrm{b}(1-\omega) \\ b_3 = (\mathrm{a}-1)(1-\omega) - \rho^2(\rho + \omega) \end{gathered}$$

so that the ∇ matrix is

$$\nabla = \begin{bmatrix} b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & b_3 & b_2 & b_1 & b_0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 \end{bmatrix}.$$

Solving this system of inequalities such that the principal minors of ∇ are greater than zero[2] yields the performance regions shown in Figure 5.8.

[2]These solutions were found numerically

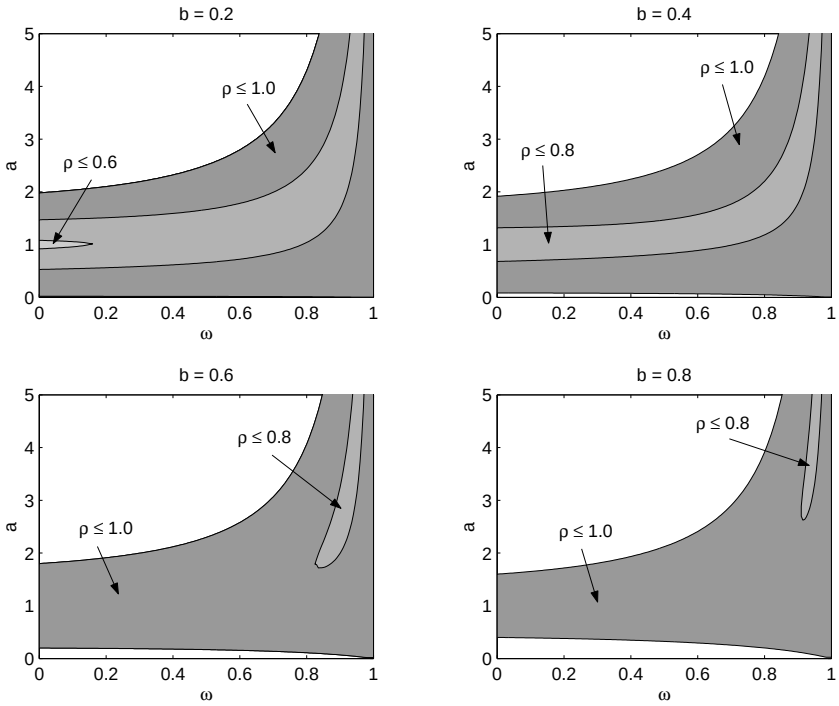

Figure 5.8: The complex performance regions for a system with a metrology delay of two runs, b = 0.2 . . . 0.8

5.2.3 Simulation of a MIMO Process

As an example of a process under MIMO EWMA control, consider the following simulation of a 4-input 2-output process with $T = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, $\nu = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$,

$$\beta = \begin{bmatrix} 6 & -3 & 8 & 16 \\ 14 & 10 & -1 & 19 \end{bmatrix}$$

and

$$b = \begin{bmatrix} 2.3 & -2.0 & 4.0 & 7.0 \\ 7.2 & 4.8 & -0.1 & 10.2 \end{bmatrix}.$$

The process has normally distributed noise of the appropriate dimension with unit variance. A MIMO EWMA controller will be used with an objective function of minimizing the 2-norm of the controller input so that

$$\boldsymbol{\xi} = \begin{bmatrix} 2.0 & 0.2 \\ -0.2 & 2.0 \end{bmatrix}$$

and the eigenvalues of $\boldsymbol{\xi}$ are $\lambda = 2.0 \pm 0.2i$. We wish to find the values of ω that will guarantee that the closed-loop eigenvalues are all inside the circle with a radius of $\rho = 0.8$. According to the performance regions shown in Figures 5.6 through 5.8, the feasible values of ω are $0.110 \leq \omega \leq 0.900$ for no delay, $0.405 \leq \omega \leq 0.911$ for a delay of one, and $0.703 \leq \omega \leq 0.921$ for a delay of two. The feasible regions for delays of up to four runs are summarized in Table 5.2.

Table 5.2 shows that a choice of $\omega = 0.90$ will achieve the desired performance for processes with delays of up to three runs. By increasing the metrology delay to four runs, however, we see that there are no values of ω that will meet this performance requirement.

The simulation is shown in Figure 5.9 where a step disturbance is introduced at run number 10. Along with the MIMO simulations, a reference performance trajectory is included which corresponds to $\lambda = 0.8$. Figure 5.9 shows that all three simulations reject the step disturbance faster than the reference trajectory which indicates that the performance requirement has been met.

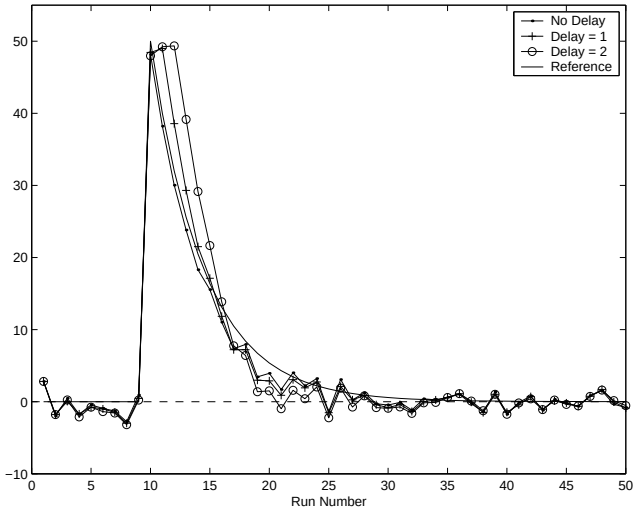

Figure 5.9: Simulation of a closed-loop MIMO system with $\lambda = 2.0 \pm 0.2i$, $\omega = 0.90$, and delays of 0, 1, and 2 runs

Table 5.2: The values of ω required to meet the performance requirement in the MIMO EWMA example for varying lengths of metrology delay

Delay	Lower Bound	Upper Bound
0	0.110	0.900
1	0.405	0.911
2	0.703	0.921
3	0.859	0.929
4	No Solution	

5.3 Conclusions

In this chapter the Routh-Hurwitz stability criterion was used to find the tuning parameters of the EWMA controller that guarantee a minimum closed-loop performance. This was accomplished by making an extension to the closed loop stability analysis of the EWMA controller used in Chapter 2. Whereas stability boundaries can be determined by mapping the unit circle to the LHP, a minimum performance boundary can be determined by mapping a circle with radius less than unity to the LHP. After this mapping step, the same Routh-Hurwitz stability criterion can be used to find the conditions where all of the closed-loop poles are inside this smaller circle. The performance regions for EWMA controlled systems with metrology delays of zero, one, and two runs were derived. This simple technique can be used to find the performance conditions for systems with larger delays.

The performance requirements for MIMO systems with several different objective functions were also derived in this chapter. When all of the eigenvalues of the model mismatch matrix are strictly real, the condition for meeting the MIMO performance requirements is that all of the eigenvalues must fall inside the SISO performance boundaries (with respect to ξ). However, unlike the SISO case, the eigenvalues of this matrix can have an imaginary component. Therefore, complex performance regions were derived using the generalized Routh-Hurwitz criteria. Several simulations were provided which show the importance in considering model mismatch and metrology delay implicitly in choosing the EWMA tuning parameters.

Chapter 6

The Transient Performance of the Double-EWMA Controller

In this chapter the performance boundaries of the double-EWMA run-to-run controller are derived. Like the performance boundaries of the EWMA controller derived in Chapter 5, the performance of the double-EMWA controller is determined by finding the conditions where all of the closed-loop poles are located inside a circle of radius $\rho < 1$ and centered at $\{0,0\}$ on the complex plane. Although the same techniques used to find the EWMA iso-performance boundaries in Chapter 5 are extended to find the performance regions of the double-EWMA controller, the characteristics of the double-EWMA performance boundaries are very different from those of the EWMA controller. One such difference is the existence of a tuning parameter that maximizes the allowable model mismatch for a given performance radius. It can be argued that this is the optimal tuning parameter for the double-EWMA controller as it corresponds to the largest allowable model mismatch for a desired closed-loop performance.

The outline of the chapter is as follows. The closed-loop transient performance requirements for the SISO double-EWMA controller are derived in Section 6.1 followed by the MIMO double-EWMA performance conditions in Section 6.2. In Section 6.3 the optimal tuning parameter for the double-EWMA controller is de-

rived. Some concluding remarks are then given in Section 6.4.

6.1 Performance of the SISO Double-EWMA Controller

It was shown in (3.19) that the characteristic equation of the closed-loop SISO double-EWMA controller is

$$f(z) = z^{d+2} + a_{d+1}z^{d+1} + a_d z^d + a_1 z + a_0 = 0 \tag{6.1}$$

where

$$\begin{aligned} a_0 &= (d(1-\omega)\omega + 1)(1-\omega)(1-\xi) \\ a_1 &= (d\omega - (d+2))(1-\omega)(1-\xi) \\ a_d &= \omega^2 \\ a_{d+1} &= -2\omega. \end{aligned}$$

Instead of simply requiring that all of the roots of (6.1) fall inside the unit circle, a minimum performance requirement can be established by solving for the conditions where the roots of (6.1) are constrained to a region smaller than the unit circle [2, 3]. By constraining the eigenvalues to this region, we can guarantee that the error decays with a rate of ρ^t, where t is the number of control steps. It is easy to see that the closer ρ is to the origin, the faster the controller error will decay.

The double-EWMA closed-loop performance boundaries will be derived for systems with metrology lags of zero, one, and two runs.

6.1.1 SISO Performance Without Metrology Delay

For a process without metrology delay, (6.1) becomes

$$f(z) = z^2 + (-2\xi\omega + 2\xi - 2)z + (\omega^2\xi - \xi + 1) = 0. \tag{6.2}$$

Applying the bilinear transformation in (5.5), which maps the performance circle to the LHP, to (6.2) yields

$$f(w) = a_2 w^2 + a_1 w + a_0 = 0 \tag{6.3}$$

where

$$\begin{aligned} a_0 &= \rho^2 - 2(1-(1-\omega)\xi)\rho - (1-\omega^2)\xi + 1 \\ a_1 &= 2(\rho^2 + (1-\omega^2)\xi - 1) \\ a_2 &= \rho^2 + 2(1-(1-\omega)\xi)\rho - (1-\omega^2)\xi + 1. \end{aligned} \tag{6.4}$$

The roots of (6.3) are all in the LHP when the coefficients on w have the same sign. Solving this system of inequalities yields three active boundaries on the closed-loop performance:

$$\begin{aligned} &1. \qquad \xi \leq \tfrac{\rho^2+2\rho+1}{-\omega^2+2\rho(1-\omega)+1} \\ &2. \qquad \xi \geq \tfrac{-1+\rho^2}{\omega^2-1} \\ &3. \qquad \xi \leq \tfrac{\rho^2-2\rho+1}{-\omega^2-2\rho(1-\omega)+1}. \end{aligned} \tag{6.5}$$

These boundaries (labelled boundaries 1 through 3) are plotted in Figure 6.1 for $\rho = 0.6$.

Figure 6.1 shows that the double-EWMA controller has a performance boundary which restricts the choice ω to $\omega \leq \rho$. This result is in contrast to the EWMA performance regions shown in Figure 6.1 where we see that closed-loop performance is not limited (strictly speaking) by the choice of ω. That is, the EWMA iso-performance regions are bounded on top and bottom but not on the right. As long as the model mismatch is large enough, *any* closed loop performance is attainable for a given ω.

This additional performance boundary (i.e. boundary 3) for the double-EWMA controller is the evolution of a boundary located at $\omega = 1$ for the stability

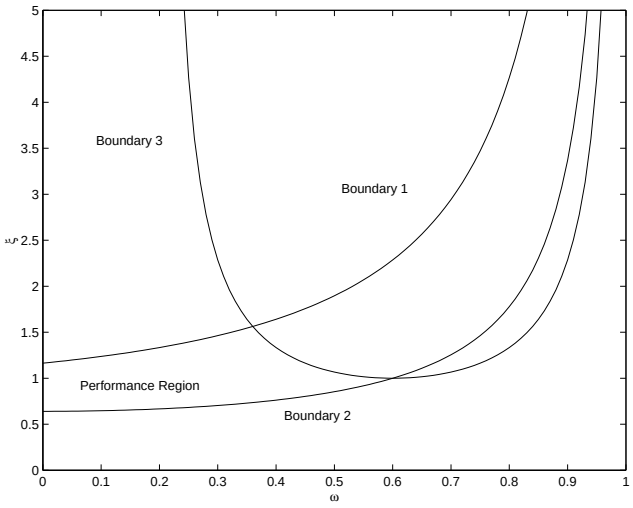

Figure 6.1: The three performance boundaries of the double-EWMA controller without metrology delay and $\rho = 0.6$

region. By reducing the performance region (i.e., decreasing ρ) this boundary moves away from $\omega = 1$ and forms the right-hand performance boundary. The EWMA stability boundary, on the other hand, does not have this boundary at $\omega = 1$ so that the right-had performance boundary never develops.

The performance boundaries for $\rho = 0.2 \ldots 1.0$ are shown in Figure 6.2. When $\rho = 1$ the performance region is equivalent to the SISO stability region so that (6.3) is the same as the closed-loop stability boundary equation in (3.5) and shown in Figure 3.1. We can clearly see in Figure 6.2 that the double-EWMA is bounded by both the tuning parameter, ω, and the model mismatch, ξ. It follows that a necessary condition for all of the roots of the double-EWMA controller falling inside the performance boundary is that the tuning parameter, ω, must be less than

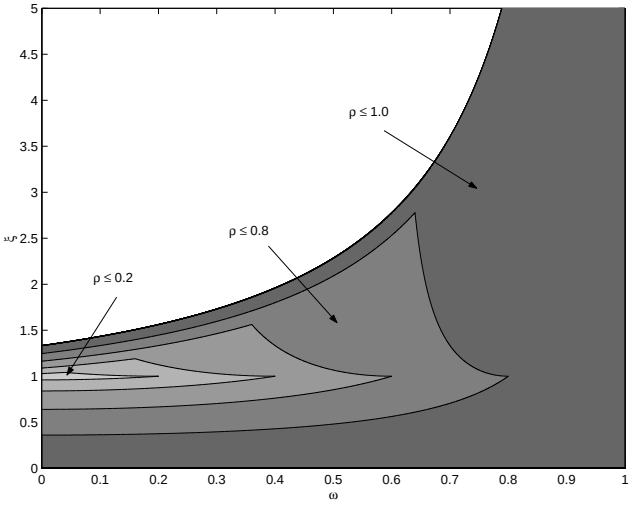

Figure 6.2: The performance region of the SISO double-EWMA system without metrology delay

or equal to the performance radius, ρ.

6.1.2 SISO Performance With Metrology Delay

6.1.2.1 A Delay of One Run

For a process with a metrology delay of one run, (6.1) becomes

$$f(z) = z^3 - 2\omega z^2 + (\omega^2\xi - 4\omega\xi + 4\omega + 3\xi - 3)z + 2(\omega\xi - \omega - \xi + 1) = 0. \quad (6.6)$$

Mapping the performance region to the LHP yields

$$f(w) = a_3w^3 + a_2w^2 + +a_1w + a_0 \quad (6.7)$$

where

$$\begin{aligned} a_0 &= \rho^3 + 2\omega\rho^2 + (\omega^2\xi - 4\omega\xi + 4\omega + 3\xi - 3)\rho - 2(\xi - 1)(1 - \omega) \\ a_1 &= 3\rho^3 - 2\omega\rho^2 - (\omega^2\xi - 4\omega\xi + 4\omega + 3\xi - 3) + 6(\xi - 1)(1 - \omega) \\ a_2 &= 3\rho^3 - 2\omega\rho^2 - (\omega^2\xi - 4\omega\xi + 4\omega + 3\xi - 3) - 6(\xi - 1)(1 - \omega) \\ a_3 &= \rho^3 + 2\omega\rho^2 + (\omega^2\xi - 4\omega\xi + 4\omega + 3\xi - 3)\rho + 2(\xi - 1)(1 - \omega). \end{aligned}$$

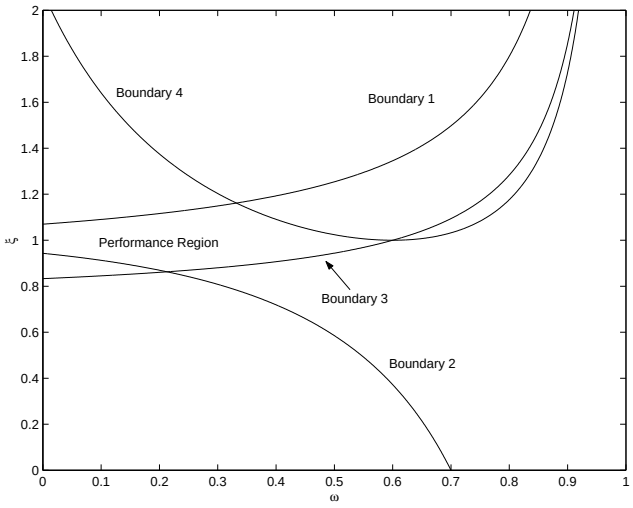

Figure 6.3: The four performance boundaries of the double-EWMA controller with a delay of one run and $\rho = 0.6$

Using the Routh-Hurwitz stability criterion to find when all of the poles of (6.7) are in the LHP yields the following four performance boundaries:

$$\begin{aligned}
&1. \quad \xi \leq 1 + \frac{\rho^4\omega+4\rho^2\omega-3\rho^4+\sqrt{\rho^6(9\rho^2-6\omega\rho^2+8\omega^2-24\omega+\rho^2\omega^2+16)}}{8(1-\omega)} \\
&2. \quad \xi \geq 1 + \frac{\rho^4\omega+4\rho^2\omega-3\rho^4-\sqrt{\rho^6(9\rho^2-6\omega\rho^2+8\omega^2-24\omega+\rho^2\omega^2+16)}}{8(1-\omega)} \\
&3. \quad \xi \geq \frac{\rho^3+2\omega\rho^2+(4\omega-3)\rho-2+2\omega}{(4\omega-\omega^2-3)\rho+2\omega-2} \\
&4. \quad \xi \leq \frac{\rho^3-2\omega\rho^2+(4\omega-3)\rho+2-2\omega}{(4\omega-\omega^2-3)\rho+2\omega-2}
\end{aligned} \tag{6.8}$$

which is equivalent to the stability criterion in (3.19) when $\rho = 1$. The performance boundaries for the case when $\rho = 0.6$ is shown in Figure 6.3.

Just as the case without metrology delay, we can see boundaries that are irrelevant for stability become active boundaries for the performance region shown in Figure 6.3. The stability region for a delay of one run (shown in Figure 3.2) has a lower stability boundary from $0 \leq \omega < 0.5$. This lower boundary explains why the

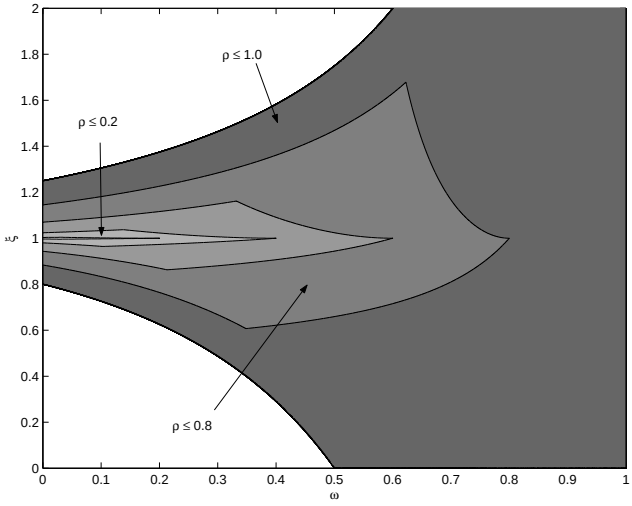

Figure 6.4: The performance region of the SISO double-EWMA system with a delay of one run

case with metrology delay has four performance boundaries as compared to three without metrology delay. The performance boundaries for $\rho = 0.2 \ldots 1.0$ are shown in Figure 6.4.

6.1.2.2 A Delay of Two Runs

By extending the delay to two runs, the characteristic equation in (6.1) becomes

$$f(z) = z^4 - 2\omega z^3 + \omega^2 z^2 + 2(\omega - 2)(\omega - 1)(1 - \xi)z - (\omega - 2)(1 - \omega)(1 - \xi) = 0 \quad (6.9)$$

and after mapping the inside of the performance region to the LHP, the characteristic equation becomes

$$f(w) = a_4 w^4 + a_3 w^3 + a_2 w^2 + a_1 w + a_0 = 0 \quad (6.10)$$

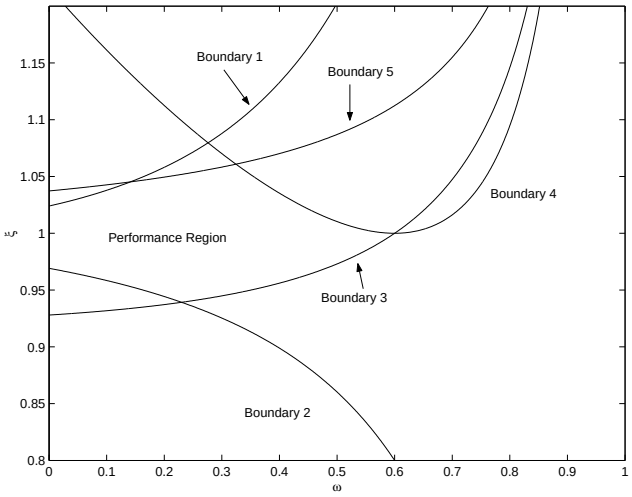

Figure 6.5: The five performance boundaries of the double-EWMA controller with a delay of two runs and $\rho = 0.6$

where

$$
\begin{aligned}
a_0 &= \rho^4 - 2\omega\rho^3 + \omega^2\rho^2 - 2(1-\omega)(\omega-2)(\xi-1)\rho + (1-\omega)(\omega-3)(\xi-1) \\
a_1 &= 4\rho^4 - 4\omega\rho^3 + 4(1-\omega)(\omega-2)(\xi-1)\rho - 4(1-\omega)(\omega-3)(\xi-1) \\
a_2 &= 6\rho^4 - 2\omega^2\rho^2 + 6(1-\omega)(\omega-3)(\xi-1) \\
a_3 &= 4\rho^4 + 4\omega\rho^3 - 4(1-\omega)(\omega-2)(\xi-1)\rho - 4(1-\omega)(\omega-3)(\xi-1) \\
a_4 &= \rho^4 + 2\omega\rho^3 + \omega^2\rho^2 + 2(1-\omega)(\omega-2)(\xi-1)\rho + (1-\omega)(\omega-3)(\xi-1).
\end{aligned}
$$

Applying the Routh-Hurwitz stability criterion on (5.12) yields five active performance boundaries. These boundaries (one of which results from the solution of a third order polynomial) are listed in Appendix C.1 and are plotted in Figure 6.5 for $\rho = 0.6$. The performance boundaries for $\rho = 0.6 \ldots 1.0$ are shown in Figure 6.6.

6.1.3 Simulation of a SISO Process

As an example, we will look at the same simulation as Section 5.1.3 where a deterministic drift of $\delta = 5$ has been added to the process and the model mismatch

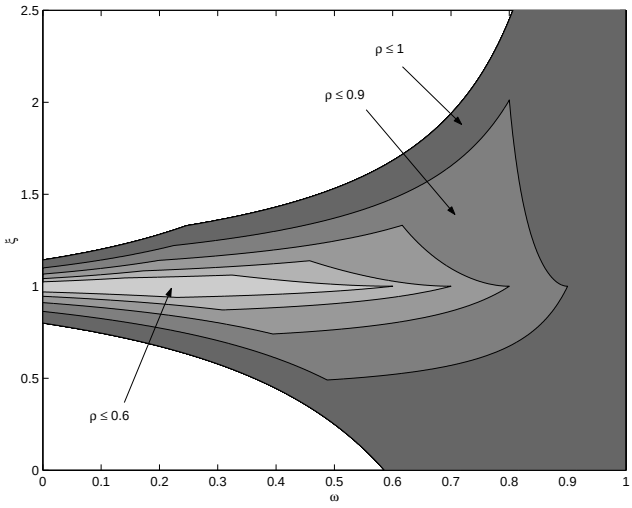

Figure 6.6: The performance region of the SISO double EWMA system with a metrology delay of two runs for $\rho = 0.6 \ldots 1$

Table 6.1: The distance from the origin to the largest closed-loop pole, ρ

Delay	$\omega = 0.2$	$\omega = 0.5$	$\omega = 0.8$
0	0.600	0.667	0.867
1	0.949	0.718	0.875
2	1.056	0.873	0.883

has been reduced to $\xi = 4/3$. Like the EWMA simulation, we will look at the step responses of processes with metrology delays of zero, one and two runs and with discounting factors of $\omega = 0.2$, 0.5, and 0.8. The results of the simulation are plotted in Figure 6.7 and the distances from the origin to the largest of the closed-loop poles are summarized in Table 6.1.

The first column of plots in Figure 6.7 shows the results of the simulation for the case without metrology delay. Here we see that increasing ω moves the

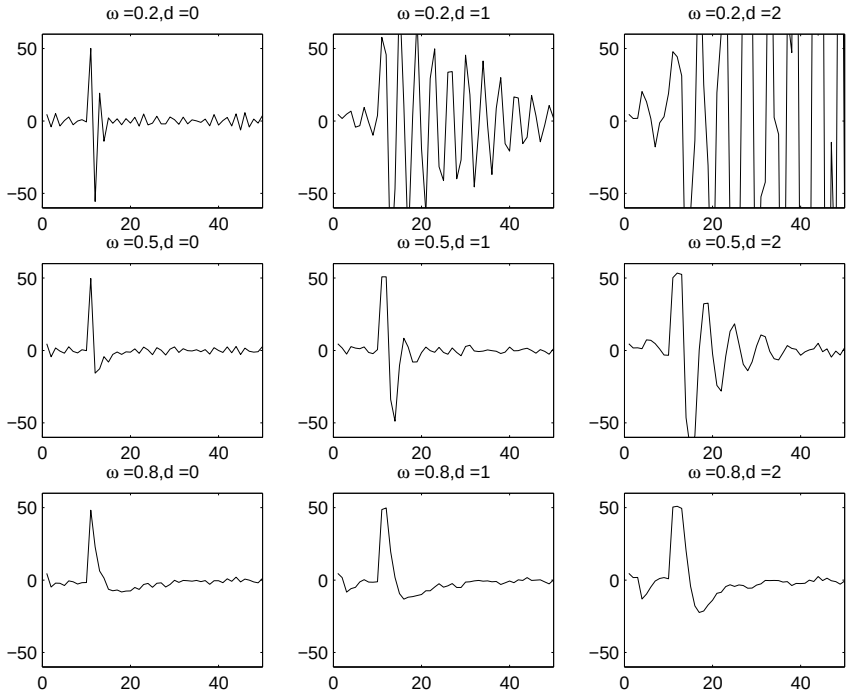

Figure 6.7: Simulations of a SISO double-EWMA controller with $\xi = 2$; $\omega = 0.2, 0.5, 0.8$; and delays of 0, 1, and 2 runs

largest closed-loop pole further from the origin, thus slowing the response to the step disturbance. Although $\omega = 0.2$ shows the fastest response to the disturbance, the controller in this case is considerably underdamped so oscillations are seen. It can be argued that these oscillations indicate a less satisfactory performance than the cases with larger ω. Nevertheless, in comparison to $\omega = 0.5$ and $\omega = 0.8$, the fastest damping occurs with $\omega = 0.2$.

The effects of metrology delay are shown in the second and third columns of Figure 6.7. Here we see that increasing the metrology delay dramatically decreases the rate at which the controller responds to the step response. A good example of this is shown in the first row of Figure 6.7. Without metrology delay, the controller shows the fastest damping to the step disturbance when $\omega = 0.2$. Then, with the addition of metrology delay, we see the closed-loop performance deteriorate. When the metrology delay is two runs the process becomes unstable.

6.2 Performance of the MIMO Double-EWMA Controller

As in Chapter 5, we can guarantee a minimum closed-loop performance for the MIMO double-EWMA controller by ensuring that all of the roots of the MIMO characteristic equation fall inside a circle with radius, ρ, on the complex plane. It follows from Theorem 3.2.2 that when the eigenvalues of the model mismatch matrix are real, then the MIMO performance requirements are equivalent to the SISO performance requirements with respect to the eigenvalues of the model mismatch matrix. The model mismatch matrices for different MIMO controller objective functions can be found in Equations (2.22) through (2.25).

When $\boldsymbol{\xi}$ has complex eigenvalues then the closed-loop characteristic equation in (6.1) has complex coefficients and the Routh-Hurwitz stability criterion is no longer valid. Nevertheless, the performance regions can be derived using the generalized Routh-Hurwitz criterion [29, 31]. This method; however, has the drawback of requiring the solution to a $4(d+1)^{th}$ order polynomial. For example, a system with a delay of only two runs requires the solution of a 12^{th} order polynomial. Therefore, when the eigenvalues of the model mismatch matrix are complex, we will take a slightly different approach for determining the performance boundaries. Starting with the MIMO double-EWMA characteristic equation

$$f(z) = z^{d+2} + a_{d+1}z^{d+1} + a_d z^d + a_1 z + a_0 = 0 \tag{6.11}$$

where

$$\begin{gathered} a_0 = (d(1-\omega)\omega + 1)(1-\omega)(1-\lambda_j) \\ a_1 = (d\omega - (d+2))(1-\omega)(1-\lambda_j) \\ a_d = \omega^2 \\ a_{d+1} = -2\omega, \end{gathered}$$

$\lambda_j = \lambda_1 \ldots \lambda_m$ are the eigenvalues of the model mismatch matrix, $\boldsymbol{\xi}$. Solving (6.11) for λ_j and changing z to polar notation, $z = \rho e^{i\theta}$, yields the complex performance boundary:

$$\lambda_j = \frac{\rho e^{(d+2)i\theta} + a_{d+1}\rho e^{(d+1)i\theta} + a_d \rho e^{i\theta d}}{a_1 \rho e^{i\theta} + a_0}. \tag{6.12}$$

This boundary will be shown for metrology delay of zero through two runs by evaluating (6.12) for $-\pi \leq \theta \leq \pi$.

6.2.1 MIMO Performance Without Metrology Delay

An example of a complex performance region for a process without metrology delay and $\rho = 0.5$ is shown in Figure 6.8. Here we see that the complex performance

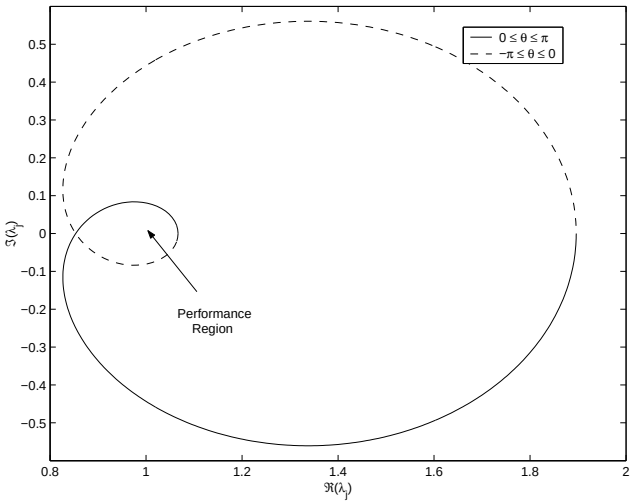

Figure 6.8: The complex (i.e., MIMO) performance boundary for a double-EWMA controller without metrology delay and $\rho = 0.5$

boundary is the region enclosed by evaluating (6.12) over the region $-0.299 \leq \theta \leq 0.299$. If all of the eigenvalues of the model mismatch matrix fall inside this region then all of the closed-loop poles will have a magnitude of less than 0.5.

The performance regions for $\omega = 0 \ldots 0.6$ (in increments of 0.2) and for $\rho = 0.65 \ldots 1.0$ (in increments of 0.05) are shown in the four plots in Figure 6.9. Here we see that as the double-EWMA tuning parameter is increased, the area of the performance region decreases. However, as ω increases, we see an increase in the size of the stability region. That is to say, there is a tradeoff between stability and performance. The bottom right plot in Figure 6.9 (corresponding to $\omega = 0.6$) shows a very small performance region for $\rho = 0.65$ (the innermost performance boundary). This is consistent with the results of the SISO double-EWMA performance analysis were we see that the maximum achievable performance occurs when $\rho = \omega$. It follows

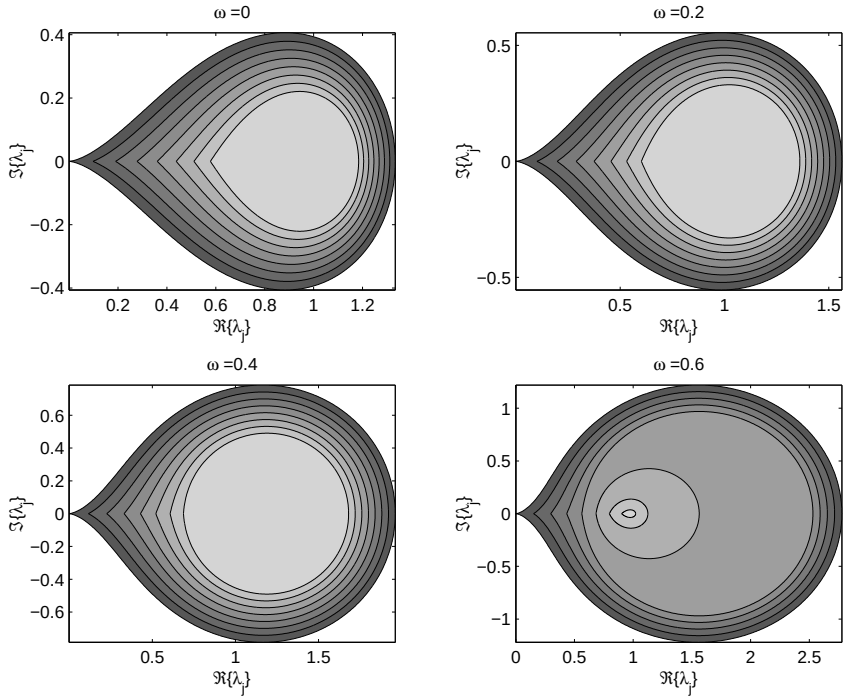

Figure 6.9: The complex performance boundaries for the double-EWMA controller without metrology delay. Shown are the boundaries for $\rho = 0.65 \ldots 1$

that $\rho \geq \omega$ is a necessary condition for achieving a desired closed-loop performance and as ω approaches ρ the area of the performance region approaches zero.

6.2.2 MIMO Performance With Metrology Delay

6.2.2.1 A Delay of One Run

The performance regions for a process with a delay of one run can be derived by evaluating (6.12) for $-\pi \leq \theta \leq \pi$ and $d = 1$. The performance regions for $\omega = 0 \ldots 0.6$ and $\rho = 0.65 \ldots 1.0$ are shown in Figure 6.10

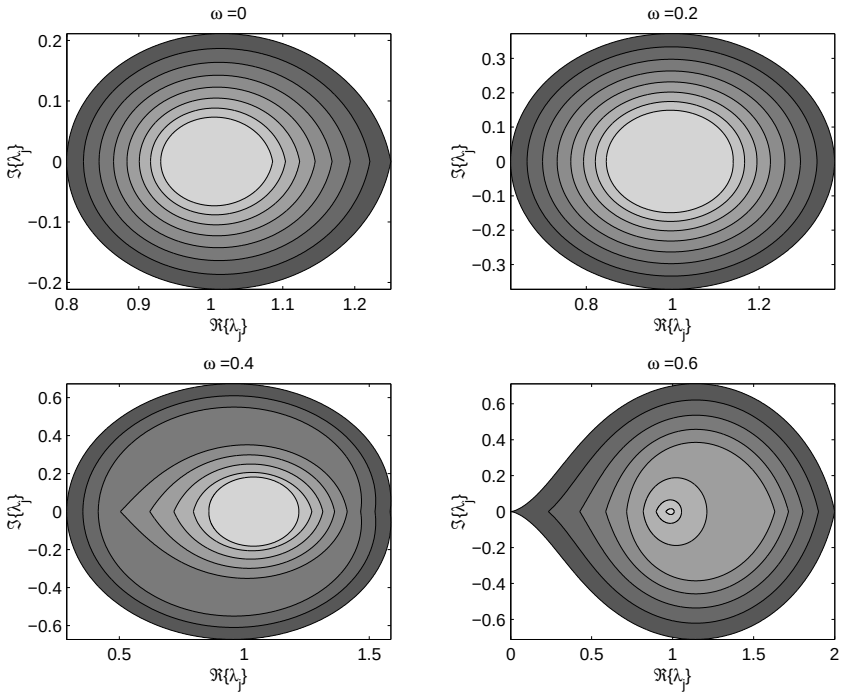

Figure 6.10: The complex performance boundaries for the double-EWMA controller with a metrology delay of one run. Shown are the boundaries for $\rho = 0.65 \ldots 1$

6.2.2.2 A Delay of Two Runs

The performance regions for the double-EWMA controller with a delay of two runs is shown in Figure 6.11. By comparing Figures 6.9 through 6.10 we see that as metrology delay increases, the size of the performance regions decreases dramatically. In the absence of metrology delay a sufficient condition for meeting the $\rho \leq 0.6$ performance requirement is that all of the eigenvalues of the model mismatch matrix fall inside a circle centered at {0.95,0} in the complex plane with a radius $r = 0.221$. With a delay of two runs, on the other hand, a sufficient condition for meeting the $\rho \leq 0.6$ performance requirement is that the eigenvalues all fall within a circle centered at {1,0} with a radius of only 0.0327. By increasing the metrology delay to only two runs, the radius of this sufficient condition for stability is reduced by 85%.

6.2.3 Simulation of a MIMO Process

As an example of a process under MIMO double-EWMA control, we will look at the same process as in Section 5.1.3 with an additional drift term of [48]

$$\delta = \begin{bmatrix} -17 \\ 1.5 \end{bmatrix}$$

and, since the double-EWMA controller is considerably less robust than the EWMA controller, the (arbitrarily chosen) estimated model gain matrix is chosen to be closer to the true process gain,

$$b = \begin{bmatrix} 3 & -4.2 & 7.2 & 11.0 \\ 13.3 & 8.3 & 0.4 & 19.3 \end{bmatrix}.$$

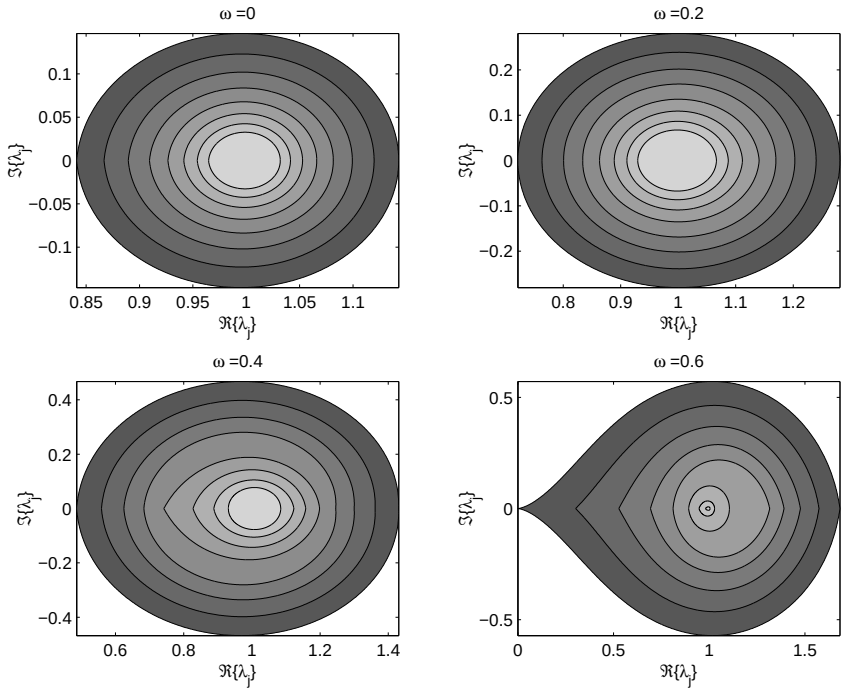

Figure 6.11: The complex performance boundaries for the double-EWMA controller with a metrology delay of two runs. Shown are the boundaries for $\rho = 0.65 \ldots 1$

Table 6.2: The values of ω required to meet the performance requirement in the MIMO double-EWMA example for varying lengths of metrology delay

Delay	Lower Bound	Upper Bound
0	0.0	0.714
1	0.191	0.682
2	0.555	0.639
3	No Solution	

The objective function for the double-EWMA controller is the same as the EWMA simulation, namely, minimize the 2-norm of the controller input so that

$$\boldsymbol{\xi} = \begin{bmatrix} 1.1 & 0.2 \\ -0.2 & 1.1 \end{bmatrix}$$

where the eigenvalues of $\boldsymbol{\xi}$ are $\lambda = 1.1 \pm 0.2i$. Like the EWMA simulation, we wish to find the values of ω that guarantee that the closed-loop poles all fall inside the circle with a radius of $\rho = 0.8$. The feasible regions for delays of up to three runs are summarized in Table 6.2.

Table 6.2 shows that a choice of $\omega = 0.60$ will achieve the desired performance for processes with delays of up to two runs. However, by increasing the metrology delay to three runs, the double-EWMA controller is incapable of achieving a closed-loop performance such that all of the poles have a magnitude less than $\rho = 0.8$ [1]. In comparing Tables 6.2 and 5.2 we see that even though the EWMA simulation has a larger model mismatch, the EWMA simulation still has a wider range of feasible choices of ω with which the closed-loop performance can be met. The double-EWMA simulation, however, includes an aggressive drift for which the EWMA controller is

[1] The minimax closed-loop pole for the system with a delay of three runs is 0.841

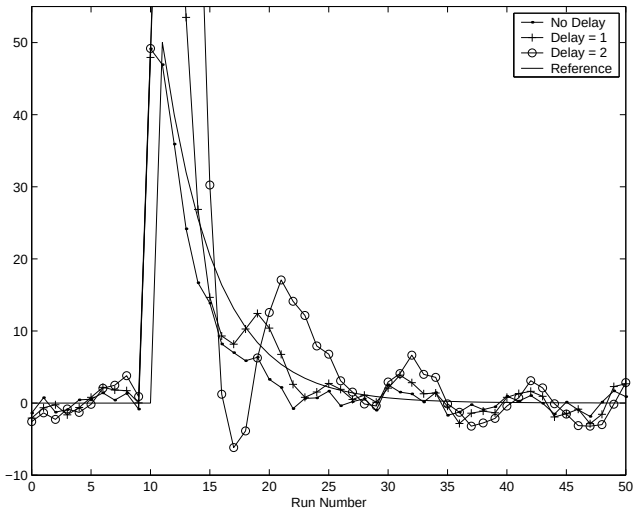

Figure 6.12: Simulations of a MIMO double-EWMA controller with $\lambda = 1.1 \pm 0.2i$; $\omega = 0.6$ and delays of 0, 1, and 2 runs

not capable of compensating. Therefore, the choice between the EWMA controller and double-EWMA controller includes a tradeoff between closed-loop performance and robustness.

The simulation is shown in Figure 6.12 where a step disturbance is introduced at run number 10. Along with the MIMO simulations, a reference performance trajectory is included which corresponds to $\lambda = 0.8$. Figure 6.12 shows that all three simulations reject the step disturbance faster than the reference trajectory which indicates that the performance requirement has been met.

6.3 The Optimal Double-EWMA Tuning Parameter

By inspecting Figures 6.1 through 6.6, we can see that a maximal model mismatch exists for a given performance radius. For example, Figure 6.2 shows that for a performance radius of $\rho = 0.6$, the iso-performance boundary peaks at $\omega = 0.36$, corresponding to a model mismatch of $\xi = 1.56$. It can be argued that $\omega = 0.36$ represents the optimal tuning parameter for the performance region of $\rho = 0.6$ as it allows for the largest model mismatch while guaranteeing a minimum closed-loop performance. In this section the optimal tuning parameter of the double-EWMA controller will be derived for systems with metrology delays of zero through two runs.

Starting with the system without metrology delay, the maximum model mismatch occurs at the intersection of the first and third boundaries in (6.5). Setting these two equations equal to one another and solving for ω yields the optimal tuning parameter

$$\omega^* = \rho^2 \tag{6.13}$$

which corresponds to a maximum model mismatch, ξ^* of

$$\xi^* = \frac{1}{1-\rho^2}. \tag{6.14}$$

This result states that if the minimum closed-loop performance is such that all of the poles fall with a radius of ρ, then a choice of $\omega = \rho^2$ will result in the maximum allowable model mismatch.

Equating the first and fourth boundaries of Equation (6.8) yields the optimal

tuning parameter for the double-EWMA controller with a delay of one run,

$$\omega^* = \rho\left(1 + \frac{\rho^2}{4} + \frac{1}{4}\sqrt{\rho^4 + 8\rho^2 - 24\rho + 16}\right). \tag{6.15}$$

The maximum model mismatch can be found by substituting (6.15) back into the first (or fourth) performance boundary in (6.8). This is provided in Appendix C.2.

The optimal tuning parameter for the case with a delay of two runs can be determined by equating the fourth and fifth boundaries in Equation (C.1). Fortunately, this problem can be simplified considerably by solving the a_2 coefficient in (6.9) for ξ,

$$\xi = \frac{3\rho^4 + 3\omega^2 - \omega^2\rho^2 - 12\omega + 9}{(1-\omega)(\omega - 3)}. \tag{6.16}$$

Although this is a valid solution to the characteristic equation in (6.9), it does not act as an active performance boundary. This is because the fifth boundary in (C.1) is closer to $\xi = 1$. However, it can be observed that (6.16) has the same intercept with the fourth boundary in (C.1) as the fifth boundary in (C.1). Therefore, the optimal tuning parameter can be determined by equating (6.16) with the fourth boundary in (C.1). It follows that the optimal performance boundary for a delay of two runs is

$$\omega^* = \frac{\gamma_1}{6(2-\rho)} + \frac{2\gamma_2}{3} - \frac{6+\rho}{3(\rho-2)} \tag{6.17}$$

where

$$\gamma_1 = \left(756\rho^5 - 2160\rho^4 + 1628\rho^3 + (1440 - 36\gamma_3)\rho^2 - (3024 - 72\gamma_3)\rho + 1728\right)^{\frac{1}{3}}$$
$$\gamma_2 = \frac{9\rho^4 - 18\rho^3 + 28\rho^2 - 42\rho + 36}{(2-\rho)\gamma_1}$$
$$\gamma_3 = -36\rho^8 + 72\rho^7 + 105\rho^6 - 252\rho^5 - 766\rho^4 + 3168\rho^3 - 4671\rho^2 + 3348\rho - 972.$$

The model mismatch evaluated at ω^* is shown in Appendix C.3.

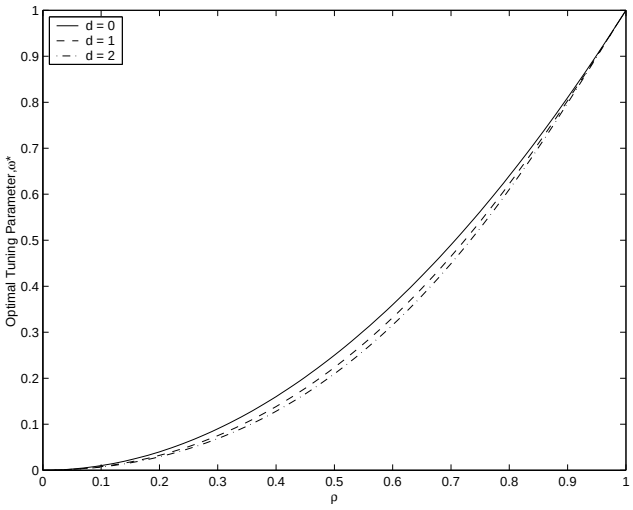

Figure 6.13: The optimal tuning parameters as a function of performance radius for delays of zero, one, and two runs

The optimal tuning parameters as a function of the performance radius are shown in Figure 6.13 for delays of zero through two runs. Here we see that the optimal tuning parameter for a performance region of radius $\rho = 0$ is $\omega = 0$. This result is intuitive in that the only way to reject a disturbance in one step is to use a dead-beat controller. We also see that the optimal tuning parameter for $\rho = 1$ (i.e. the stability boundary) is $\omega = 1$. This is equivalent to turning the controller off as the model parameters are never updated. Finally, we see in Figure 6.13 the somewhat surprising result that the optimal tuning parameters are very similar for the three delay scenarios. This suggests that the expression for the optimal tuning parameter for the system without metrology delay in (6.13) offers a satisfactory approximation for the optimal tuning parameters for systems with extended metrology delays.

6.4 Conclusions

In this chapter the requirements on the closed-loop performance of the double-EWMA controller were derived by constraining the closed-loop poles to a circle centered at the origin of the complex plane and with a radius of smaller than unity. By guaranteeing that all of the poles are inside this performance region, a minimum closed-loop performance can be chosen and an appropriate tuning parameter can be selected based on the estimated error in the process gain matrix.

Unlike the EWMA controller shown in the previous chapter, the double-EWMA controller shows a maximum performance for a given tuning parameter. More specifically, the highest attainable performance is $\omega = \rho$. This result shows that there is a tradeoff between performance and robustness in tuning the double-EWMA controller. By increasing the tuning parameter, the area of the stabile region increases, however, increasing the tuning parameter places a limit on the maximum attainable closed-loop performance.

Finally, the optimal double-EWMA tuning parameters as a function of the performance radius were derived. The optimal tuning parameters were chosen such that the system has the largest possible model mismatch while meeting the closed-loop performance requirement.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

There have been many studies on the closed loop stability and performance of run-to-run controllers to date (for instance [2, 3, 14, 17, 37, 43, 63]). Although the majority of run-to-run controlled processes include metrology delay, the effect of metrology delay on the stability and performance run-to-run controlled processes has been largely ignored (with the notable exception of the work by Adivikolnu and Zafiriou where the a process with a delay of one run was studied [2, 3]). In addition, the published work to date on the stability and performance of MIMO run-to-run controllers has been limited to a small subset of the published MIMO control laws. The broad goal of this work has been to derive the conditions for closed-loop stability and performance of the EWMA and double-EWMA run-to-run controllers with metrology delay for SISO systems and for a multiplicity of MIMO control laws.

A summary of the results presented in this dissertation are as follows:

- By expressing the closed-loop realization of run-to-run controlled processes as a function of the model mismatch and metrology delay, well understood stability analysis techniques (e.g., the Routh-Hurwitz stability criterion) can be used

to find the stability conditions for the SISO EWMA and double-EWMA run-to-run controllers. The main result of this analysis, although admittedly not surprising, shows that the gain margin shrinks as the metrology delay increases and as the tuning parameter approaches zero.

- The eigenvalues of the model mismatch matrix determines the stability of the MIMO EWMA and double-EWMA controllers. When the eigenvalues of the model mismatch matrix of the MIMO system are all real, the necessary and sufficient condition for stability is that all of the eigenvalues fall inside the stability regions derived in the SISO analysis. When the eigenvalues are complex, the Routh-Hurwitz stability criterion is no longer valid. The generalized Routh-Hurwitz stability criterion, however, can be used to find analytical expressions for the stability of the MIMO EWMA and double-EWMA controllers.

- A sufficient condition for MIMO EWMA stability for any length metrology delay is that all of the eigenvalues of the model mismatch matrix fall inside a circle centered at {1,0} on the complex plane and with unit radius. It follows that a sufficient condition for stability of the SISO EWMA controller (which is simply a subset of the MIMO result) is that the ratio of true process gain and the estimated process gain must be between zero and two for any length of metrology delay.

- A sufficient condition for stability of the MIMO double-EWMA controller is that all of the eigenvalues of the model mismatch matrix fall inside a circle

centered at {1,0} on the complex plane and with a radius of

$$r = \frac{d}{2d+3}.$$

A sufficient condition for stability of the SISO double-EWMA controller is that the ratio of the true process gain and the estimated process gain is between $1 - \frac{d}{2d+3}$ and $1 + \frac{d}{2d+3}$. A necessary condition for closed-loop stability was also derived. The size of this necessary condition for stability approaches zero as the metrology delay approaches infinity indicating that any error between the estimated process model and the true process model will result in closed-loop instability.

- When the feasible region of the EWMA and double-EWMA tuning parameters are mapped to the right half plane, a nearly linear relationship is observed with the stability boundaries for both SISO and MIMO systems. This relationship can be used to generate accurate numerical approximations of the stability boundaries and these approximations can be found in Chapter 4. In addition, these relationships can be used to derive a sufficient condition and a necessary condition for closed-loop stability.

- Minimum performance boundaries can be derived by solving for the conditions where all of the closed-loop poles fall inside a circle centered at the origin and with a radius smaller than the unit circle. The techniques used in the stability analysis can be extended to the performance analysis by using a bilinear transformation to map the performance region to the open left half plane. The Routh-Hurwitz (and generalized Routh-Hurwitz) stability criterion can then

be used to find the performance boundaries of the EWMA and double-EWMA controllers.

- Finally, using the performance analysis procedure, an optimal double-EWMA tuning parameter can be derived for a desired minimum performance requirement. This tuning parameter allows for the largest model mismatch while guaranteeing that a minimum performance requirement (in terms of pole placement) is met. While this optimal tuning parameter is complicated for systems with metrology delay, the optimal tuning parameter can be approximated with the optimal tuning parameter of a system without metrology delay. A satisfactory choice for double-EWMA tuning parameter is $\omega = \rho^2$ regardless of the length of metrology delay.

7.2 Recommendations for Future Work

This work has presented an analysis of the closed-loop stability and performance of semiconductor processes controlled by EWMA and double-EWMA controllers. However, as in most research projects, the in-depth study of a problem often poses more questions that it solves. Some of the areas for further research in the stability and performance of run-to-run controllers are discussed in this section.

7.2.1 Variable Metrology Delay Lengths

The entirety of the work in this dissertation is based around the assumption that the delay between product manufacturing and product metrology is a constant. This, however, is rarely the case in semiconductor manufacturing as fluctuations in

tool utilization results in variable metrology delay lengths. The following topics are in need of further investigation:

- Some preliminary results have indicated that systems with normally distributed metrology delay are more stable than predicted using the results of this work in respect to the mean metrology delay. The existence of shorter metrology delays in a mixed delay system appears to have a stabilizing effect on the overall process. It follows that the distribution of the metrology delay as well as the mean metrology delay has an effect on the stability of the process. The effect of metrology delay distribution on run-to-run stability remains very much an open topic.

- In addition to metrology delay distribution, the effect of queuing discipline (e.g., first in first out (FIFO) or last in first out (LIFO)) and out of order metrology on system stability remain as open topics. Preliminary results have shown that the LIFO queueing discipline results in greater robustness to model mismatch than FIFO systems. Again, we see that the existence of some runs with shorter metrology delays in LIFO queuing systems results in greater robustness to model mismatch than the pure delay systems considered in this work.

- The use of seasonal indicator models to estimate the disturbance contributions from a multiplicity of sources (e.g., tool, product, reticle, etc.) has recently received considerable attention [27, 47, 51]. In this run-to-run framework, there exists the potential for a different metrology delay length for each of the dis-

turbance channels. The effects of model mismatch on the stability of such a run-to-run system remains an open topic.

7.2.2 Asymptotic Variance Estimation

It was mentioned in Section 1.1.3 that performance analysis of run-to-run controllers have focused on both transient behavior and long-term behavior. However, the closed-loop performance analysis shown in Chapters 5 and 6 have dealt exclusively with the effect of metrology delay on the *transient* performance of run-to-run controllers. The effect of metrology delay, model mismatch, tuning parameter, and ARIMA parameters on the long term variance of run-to-run controllers remains as an open topic and further investigation in this field is necessary.

7.2.3 The Stability of Other Run-to-Run Controllers

This work has focused on the stability of the EWMA and double-EWMA run-to-run controllers. Although these two algorithms represent the two most frequently implemented controllers in industry, many additional run-to-run controllers algorithms have been proposed (see Section 1.1.1.3). It follows that a study of the stability and performance of these controllers would be beneficial and remains, for the most part, as an open topic.

In addition to investigating the stability and performance of other run-to-run controllers, many details of the stability and performance of the EWMA and double-EWMA controllers remain as open topics. For instance, throughout this work, it has been assumed that the underlying process models have a constant gain and a

drifting disturbance. Although this type of model forms the basis of the majority of run-to-run controllers, there are several control scenarios where the run-to-run filters are used to recursively estimate the process gain while the process intercept remains constant (for example CMP polish rate [58, 61, 68] and etch rate [20]). The stability and performance of these systems are not considered in this work and this topic remains open for further investigation.

7.2.4 Closed Loop Performance Monitoring

Throughout this work the stability and performance of a closed-loop process has been derived for a given model mismatch and controller tuning parameter. The question remains as to whether closed-loop performance (through disturbance rejection characteristics or closed-loop system identification) can be used to estimate model mismatch. Finally, by periodically monitoring and estimating the closed loop performance, it may be possible to obtain more accurate estimates of the process gain and continually improve the stability and performance of run-to-run controllers.

Appendices

Appendix A

A.1 Derivation of MIMO Model Mismatch Matrices

A.1.1 Model Mismatch Matrix for (2.22)

Starting with the updating expression for the EWMA disturbance estimate in (1.3) and substituting the controller input in (2.22),

$$u_t = -b^T(bb^T)^{-1}\hat{\nu}_t,$$

and the true process model,

$$y_t = \beta u_t + \nu_t,$$

yields

$$\hat{\nu}_{t+1} = (I - (I - \boldsymbol{\omega})\boldsymbol{\xi})\hat{\nu}_t + (I - \boldsymbol{\omega})\nu_t, \tag{A.1}$$

where $\boldsymbol{\xi} = \beta b^T(bb^T)^{-1}$ and where the process target, T, has been set to zero without loss of generality.

A.1.2 Model Mismatch Matrix for (2.23)

Rearranging the controller input, in (2.23) yields

$$\Delta u_t = -b^T(bb^T)^{-1}\hat{v}_t - b^T(bb^T)^{-1}bu_{t-1} \tag{A.2}$$

where Δu_t is change in recipe from one run to the next, $\Delta u_t = u_t - u_{t-1}$. Next, substituting

$$u_t = -b^T(bb^T)^{-1}\hat{\nu}_{t-1} + (I - b^T(bb^T)^{-1}b)u_{t-1} \tag{A.3}$$

back into (A.2) yields

$$\Delta u_{t+1} = -b^T(bb^T)^{-1}\Delta\hat{v}_t. \tag{A.4}$$

Since the matrix, $b^T(bb^T)^{-1}b$ is idempotent, it follows that the updating equations for the change in the EWMA disturbance estimate is

$$\Delta\hat{\nu}_{t+1} = (I - (I - \boldsymbol{\omega})\boldsymbol{\xi})\Delta\hat{\nu}_t + (I - \boldsymbol{\omega})\Delta\nu_t \tag{A.5}$$

where $\boldsymbol{\xi} = \beta b^T(bb^T)^{-1}$ and where the process target, T, has been set to zero without loss of generality. It is easy to see that (A.5) is equivalent to (A.1) with respect to the change in the disturbance estimate, $\Delta\hat{\nu}$.

A.1.3 Model Mismatch Matrix for (2.24)

Starting with the updating expression for the EWMA disturbance estimate in (1.3) and substituting the controller input in (2.24),

$$u_t = -(b^Tb)^{-1}b^T\hat{\nu}_t,$$

and the true process model,

$$y_t = \beta u_t + \nu_t$$

yields,

$$\hat{\nu}_{t+1} = (I - (I - \boldsymbol{\omega})\boldsymbol{\xi})\hat{\nu}_t + (I - \boldsymbol{\omega})\nu_t \tag{A.6}$$

where $\boldsymbol{\xi} = \beta(b^Tb)^{-1}b^T$ and where the process target, T, has been set to zero without loss of generality.

A.1.4 Model Mismatch Matrix for (2.25)

Setting $S = 0$ and substituting the controller input in (2.25),

$$u_t = -(b^T Q b + R + S)^{-1}(S u_{t-1} + b^T Q \hat{\nu}_t)$$

into the true process model and the updating equation for the EWMA disturbance estimate yields

$$\hat{\nu}_{t+1} = \hat{\nu}_t + (1-\omega)\left(-\beta(b^T Q b + R)^{-1} b^T Q \hat{\nu} + b(b^T Q b + R)^{-1} b^T Q \hat{\nu}_t - \hat{\nu}_t + \nu_t\right) \quad \text{(A.7)}$$

so that rearranging yields

$$\hat{\nu}_{t+1} = (I - (I - \boldsymbol{\omega})\boldsymbol{\xi})\hat{\nu}_t + (I - \boldsymbol{\omega})\nu_t \quad \text{(A.8)}$$

where

$$\boldsymbol{\xi} = (\beta - b)(b^T Q b + R)^{-1} b^T Q + I$$

and where the process target, T, has been set to zero without loss of generality.

A.2 Proof of Lemma 2.2.1

Consider the transition matrix shown in (2.28) with the model mismatch matrix, $\boldsymbol{\xi}$ decomposed using the similarity transform, $\boldsymbol{\xi} = PDP^{-1}$ where P is a square matrix containing the eigenvectors of $\boldsymbol{\xi}$ and D is the Jordan form of $\boldsymbol{\xi}$. Since $\boldsymbol{\omega} = \omega I$,

$$A = P(I - D(I - \omega))P^{-1}. \quad \text{(A.9)}$$

Therefore,

$$eig(A) = eig(I - D(I - \omega)) = 1 - \lambda_j(1 - \omega). \quad \text{(A.10)}$$

A.3 Proof of Theorem 2.2.2

The necessary and sufficient condition for stability of the closed loop MIMO EWMA system is that all of the eigenvalues of the transition matrix lie inside of the unit circle. From Lemma 2.2.1, the eigenvalues of the MIMO transition matrix in (2.28) can be written

$$eig(\mathbf{A}) = eig \begin{bmatrix} 1-\lambda_1(1-\omega) & & \\ & \ddots & \\ & & 1-\lambda_m(1-\omega) \end{bmatrix} \tag{A.11}$$

where $\lambda_1 \ldots \lambda_m$ are the eigenvalues of $\boldsymbol{\xi}$. By comparing (A.11) with the SISO characteristic equation, (2.9), it is easy to see that the diagonal elements of $\mathbf{A}$ are equivalent to the SISO transition matrix with the SISO plant-model mismatch, ξ, replaced with the eigenvalues of the MIMO plant-model mismatch matrix, $\boldsymbol{\xi}$. Therefore, the MIMO system is stable when $0 < \lambda_j < \frac{2}{1-\omega}$ for all λ_j.

A.4 Proof of Theorem 2.2.3

The characteristic equation for $\omega = 0$ is

$$z^{d+1} = 1 - \mathrm{a} - \mathrm{b}i. \tag{A.12}$$

To require the roots $|z| < 1$ is equivalent to requiring $|1 - \mathrm{a} - \mathrm{b}i| < 1$ or

$$(1-\mathrm{a})^2 + \mathrm{b}^2 < 1. \tag{A.13}$$

Appendix B

B.1 Proof of Lemma 3.2.1

Consider the stability matrix shown in (3.26) with the plant model mismatch matrix, $\boldsymbol{\xi}$, decomposed into the similarity transform, $\boldsymbol{\xi} = PDP^{-1}$ where P is a square matrix containing the eigenvectors of $\boldsymbol{\xi}$ and D is the Jordan form of $\boldsymbol{\xi}$,

$$A = \begin{bmatrix} I_m - (I_m - \boldsymbol{\omega}^2)PDP^{-1} & I_m - (I_m - \boldsymbol{\omega}^2)PDP^{-1} \\ (I_m - \boldsymbol{\omega})^2 PDP^{-1} & I_m - (I_m - \boldsymbol{\omega})^2 PDP^{-1} \end{bmatrix}. \tag{B.1}$$

Since $eig(AB) = eig(BA)$, (B.1) can be written

$$eig(A) = eig\left(P^{-1} \begin{bmatrix} P - (I_m - \boldsymbol{\omega}^2)PD & P - (I_m - \boldsymbol{\omega}^2)PD \\ (I_m - \boldsymbol{\omega})^2 PD & P - (I_m - \boldsymbol{\omega})^2 PD \end{bmatrix} \right) \tag{B.2}$$

and since $\boldsymbol{\omega} = \omega I$, (B.2) can be rearranged to

$$eig(A) = eig\left(\begin{bmatrix} I_m - (I_m - \boldsymbol{\omega}^2)D & I_m - (I_m - \boldsymbol{\omega}^2)D \\ (I_m - \boldsymbol{\omega})^2 D & I_m - (I_m - \boldsymbol{\omega})^2 D \end{bmatrix} \right) \tag{B.3}$$

so that the eigenvalues of $\boldsymbol{\xi}$ can be considered in the stability analysis of the MIMO double EWMA controller instead of $\boldsymbol{\xi}$ itself. This greatly reduces the complexity in determining the MIMO stability region.

B.2 Proof of Theorem 3.2.2

Since D is a diagonal matrix, (B.3) can be written

$$eig(A) = eig\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right) \tag{B.4}$$

where

$$A_{11} = A_{12} = \begin{bmatrix} 1-(1-\omega^2)\lambda_1 & & \\ & \ddots & \\ & & 1-(1-\omega^2)\lambda_j \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} (1-\omega)^2\lambda_1 & & \\ & \ddots & \\ & & (1-\omega)^2\lambda_m \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} 1-(1-\omega)^2\lambda_1 & & \\ & \ddots & \\ & & 1-(1-\omega)^2\lambda_m \end{bmatrix}$$

which after elementary row and column operations can be written

$$eig(A) = eig \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{bmatrix} \tag{B.5}$$

where

$$A_1 = \begin{bmatrix} 1-(1-\omega^2)\lambda_1 & 1-(1-\omega^2)\lambda_1 \\ -(1-\omega)^2\lambda_1 & 1-(1-\omega)^2\lambda_1 \end{bmatrix}$$

$$\vdots$$

$$A_m = \begin{bmatrix} 1-(1-\omega^2)\lambda_m & 1-(1-\omega^2)\lambda_m \\ -(1-\omega)^2\lambda_m & 1-(1-\omega)^2\lambda_m \end{bmatrix}$$

where $\lambda_1 \ldots \lambda_m$ are the eigenvalues of $\boldsymbol{\xi}$.

B.3 Correction to Previous Work

Del Castillo and Rajagopal [15, 17] have recently published an analysis on the stability boundaries of the PCC. The transition matrix of this particular double-EWMA controller is

$$A = \begin{bmatrix} I-(I-\omega_1)\boldsymbol{\xi} & (I-\omega_1)(I-\boldsymbol{\xi}) \\ -(I-\omega_2)\boldsymbol{\xi} & I-(I-\omega_2)\boldsymbol{\xi} \end{bmatrix} \tag{B.6}$$

and the stability conditions for this process were published as follows:

$$|1 - 0.5\xi_{jj}(2 - \omega_1 - \omega_2) + 0.5z| < 1, \forall j \in \{1, 2 \dots m\} \tag{B.7}$$

$$|1 - 0.5\xi_{jj}(2 - \omega_1 - \omega_2) - 0.5z| < 1, \forall j \in \{1, 2 \dots m\} \tag{B.8}$$

where $z = \sqrt{\xi_{jj}(2 - \omega_1 - \omega_2)^2 - 4(1 - \omega_1)(1 - \omega_2)\xi_{jj}}$ and ξ_{jj} is the j^{th} diagonal element of the model mismatch matrix, $\boldsymbol{\xi}$. These stability conditions imply that the stability of the transition matrix in (B.6) is independent of the off-diagonal elements of $\boldsymbol{\xi}$. The following example, however, will demonstrate the eigenvalues of the transition matrix, and therefore the stability of the system, are in fact dependent on the off-diagonal terms of $\boldsymbol{\xi}$. Consider the case when the plant gain is

$$\beta = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

and the estimated gain is

$$b = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so that the model mismatch matrix is

$$\boldsymbol{\xi} = \beta b^{-1} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

We have also chosen the double-EWMA tuning parameters to be $\omega_1 = \omega_2 = 0.5$. According to the stability region in (B.7) and (B.8), this model mismatch and tuning parameter combination will result in a stable system. This however, is not the case due to the off diagonal terms of $\boldsymbol{\xi}$. When the off-diagonal elements of $\boldsymbol{\xi}$ are removed, the eigenvalues of the resulting transition matrix, A' are $eig(A') = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}$. However, the eigenvalues of the transition

matrix when the off diagonal elements of $\boldsymbol{\xi}$ are included in (B.6) are $eig(A) = \begin{bmatrix} 2.21 & 0.79 & 0.72 & -1.72 \end{bmatrix}$ where the presence of eigenvalues outside the unit circle indicate that this is in fact an unstable system. The off-diagonal elements of $\boldsymbol{\xi}$ clearly have an affect on the stability of the closed loop process.

Similar derivations to those shown in Lemma 3.2.1 can be used to show that the stability region in (B.7) and (B.8) is valid when the diagonal elements of $\boldsymbol{\xi}$ are replaced with the *eigenvalues* of $\boldsymbol{\xi}$. Returning to the example, we see that the eigenvalues of $\boldsymbol{\xi}$ are $\lambda = \{-1, 3\}$ which are outside of the stability region defined by (B.7) and (B.8).

B.4 Proof of Theorem 3.2.3

If the smallest radius from $\lambda_j = 1$ to the stability boundary is

$$r = \frac{1}{2d+3}$$

then it follows that there are no points inside of the circle centered at {1,0} on the complex plane with radius, r that are unstable. Therefore, a sufficient condition for stability is

$$(1-\mathrm{a})^2 + \mathrm{b}^2 < \frac{1}{2d+3}.$$

Appendix C

C.1 Double-EWMA Performance Boundaries for $d = 2$

Four of the five performance boundaries for the double-EWMA controller with a delay of two runs are as follows:

$$\begin{array}{ll} 1. & \xi \leq -\frac{1}{3}\frac{3\rho^4+3\omega^2-\omega^2\rho^2-12\omega+9}{(1-\omega)(\omega-3)} \\ 2. & \xi \geq \frac{(\rho+1)(\rho^3+\rho^2\omega-\rho^2+\rho-\rho\omega-3+4\omega-\omega^2)}{(1-\omega)(\rho\omega+\omega-2\ \rho-3)} \\ 3. & \xi \geq -\frac{(\rho-1)(\rho^3-\rho^2\omega+\rho^2-\rho\omega+\rho+3-4\omega+\omega^2)}{(1-\omega)(\rho\omega-\omega+3-2\rho)} \\ 4. & \xi \leq \frac{(\rho-1)^2(\rho^2-2\rho\omega+2\rho+3-4\omega+\omega^2)}{(1-\omega)(-\omega+2\rho\omega+3-4\rho)}. \end{array} \tag{C.1}$$

The fifth performance boundary can be derived by defining

$$\begin{aligned} \gamma_{41} &= (1-\omega)(\omega+2\rho\omega-4\rho-3) \\ \gamma_{40} &= (\rho+1)^2(\rho^2+2\rho\omega-2\rho+3+\omega^2-4\omega) \end{aligned}$$

$$\begin{aligned} \gamma_{31} &= -4(1-\omega)(\rho\omega+\omega-3-2\rho) \\ \gamma_{30} &= -4(\rho+1)(-\rho^3-\rho^2\omega+\rho^2+\rho\omega-\rho+\omega^2+3-4\omega) \end{aligned}$$

$$\begin{aligned} \gamma_{21} &= 6(1-\omega)(\omega-3) \\ \gamma_{20} &= 6\rho^4-2\omega^2\rho^2+18+6\omega^2-24\omega \end{aligned}$$

$$\begin{aligned} \gamma_{11} &= 4(1-\omega)(\rho\omega-\omega-2\rho+3) \\ \gamma_{10} &= 4(\rho-1)(\rho^3-\rho^2\omega+\rho^2-\rho\omega+\rho-4\omega+3+\omega^2) \end{aligned}$$

$$\begin{aligned} \gamma_{01} &= -(1-\omega)(2\rho\omega-\omega-4\rho+3) \\ \gamma_{00} &= (\rho-1)^2(\rho^2-2\rho\omega+2\rho+3-4\omega+\omega^2) \end{aligned}$$

and setting the derivative of the third principal minor equal to zero,

$$H = \left| \begin{bmatrix} \gamma_{31}\xi+\gamma_{30} & \gamma_{11}\xi+\gamma_{10} & 0 \\ \gamma_{41}\xi+\gamma_{40} & \gamma_{21}\xi+\gamma_{20} & \gamma_{01}\xi+\gamma_{00} \\ 0 & \gamma_{31}\xi+\gamma_{30} & \gamma_{11}\xi+\gamma_{10} \end{bmatrix} \right| = 0$$

so that

$$0 = \psi_3\xi^3 + \psi_2\xi^2 + \psi_1\xi + \psi_0 \tag{C.2}$$

where

$$\begin{gathered}\psi_3 = \gamma_{31}\gamma_{21}\gamma_{11} - \gamma_{41}\gamma_{11}^2 - \gamma_{01}\gamma_{31}^2 \\ \psi_2 = \gamma_{31}\gamma_{21}\gamma_{10} + \gamma_{30}\gamma_{21}\gamma_{11} - \gamma_{40}\gamma_{11}^2 - 2\gamma_{41}\gamma_{11}\gamma_{10} - 2\gamma_{31}\gamma_{01}\gamma_{30} - \gamma_{31}^2\gamma_{00} + \gamma_{31}\gamma_{20}\gamma_{11} \\ \psi_1 = -2\gamma_{40}\gamma_{11}\gamma_{10} + \gamma_{30}\gamma_{21}\gamma_{10} + \gamma_{30}\gamma_{20}\gamma_{11} - \gamma_{41}\gamma_{10}^2 - 2\gamma_{31}\gamma_{00}\gamma_{30} - \gamma_{01}\gamma_{30}^2 + \gamma_{31}\gamma_{20}\gamma_{10} \\ \psi_0 = \gamma_{30}\gamma_{20}\gamma_{10} - \gamma_{00}\gamma_{30}^2 - \gamma_{40}\gamma_{10}^2.\end{gathered}$$

Solving (C.2) for ξ gives,

$$\xi \leq -\frac{1}{12}\frac{\alpha^{\frac{1}{3}}}{\psi_3} + \frac{1}{3}\frac{3\psi_1\psi_3 - \psi_2^2}{\psi_3\alpha^{\frac{1}{3}}} - \frac{\psi_2}{3\psi_3} - \frac{i\sqrt{3}}{2}\left(\frac{\alpha^{\frac{1}{3}}}{6\psi_3} + \frac{2}{3}\frac{3\psi_1\psi_3 - \psi_2^2}{\psi_3\alpha^{\frac{1}{3}}}\right), \tag{C.3}$$

where

$$\begin{aligned}\alpha = {} & 36\psi_1\psi_2\psi_3 - 108\psi_0\psi_3^2 - 8\psi_2^3 + \\ & 12\sqrt{3}(4\psi_1^3\psi_3 - \psi_1^2\psi_2^2 - 18\psi_0\psi_1\psi_2\psi_3 + 27\psi_0^2\psi_3^2 + 4\psi_0\psi_2^3)^{\frac{1}{2}}\psi_3.\end{aligned}$$

C.2 The Maximum Model Mismatch for $d = 1$

Substituting (6.15) back into the first performance boundary in (6.8) yields

$$\xi^* = \frac{\frac{\rho^5}{2} - \rho^4 - (4+\gamma)\rho^2 - \left(3 + \frac{3\gamma}{2}\right)\rho + 6}{\frac{\rho^7}{16} + \left(\frac{1}{2} + \frac{\gamma}{8}\right)\rho^5 - \rho^4 - \left(\frac{3}{2} - \left(1 + \frac{\gamma}{4}\right)^2\right)\rho^3 - (4+\gamma)\rho^2 - \left(3 + \frac{3\gamma}{2}\right)\rho + 6} \tag{C.4}$$

where

$$\gamma = \sqrt{\rho^4 + 8\rho^2 - 24\rho + 16}.$$

C.3 The Maximum Model Mismatch for $d = 2$

Substituting (6.17) back into (6.16) yields

$$\xi^* = \frac{1}{3}\frac{3\rho^4 - \gamma_4^2\rho^2 + 3\gamma_4^2 - \frac{2\gamma_2^{\frac{1}{3}}}{\rho-2} - 8\gamma_3 + 4\frac{6+\rho}{\rho-2}}{3 - \frac{2\gamma_2^{\frac{1}{3}}}{3(\rho-2)} - \frac{8\gamma_3}{3} + \frac{4(6+\rho)}{3(\rho-2)} + \gamma_4^2} \tag{C.5}$$

where

$$\gamma_1 = -36\rho^8 + 72\rho^7 + 105\rho^6 - 252\rho^5 - 766\rho^4 + 3168\rho^3 - 4671\rho^2 + 3348\rho - 972$$
$$\gamma_2 = -756\rho^5 + 2160\rho^4 - 1628\rho^3 + (36\sqrt{\gamma_1} - 1440)\rho^2 - (72\sqrt{\gamma_1} - 3024)\rho - 1728$$
$$\gamma_3 = \frac{9\rho^4 - 18\rho^3 + 28\rho^2 - 42\rho + 36}{(\rho - 2)\gamma_2^{\frac{1}{3}}}$$
$$\gamma_4 = \frac{\gamma_2^{\frac{1}{3}} - 2\rho - 12}{\rho - 2} + \frac{2\gamma_3}{3}.$$

Bibliography

[1] B. Abraham and J. Ledolter. *Statistical Methods for Forecasting*. John Wiley and Sons, 1983.

[2] S. Adivikolanu and E. Zafiriou. Robust run-to-run control for semiconductor manufacturing: an internal model control approach. *Proceedings of the American Control Conference*, 6:3687–3691, 1998.

[3] S. Adivikolanu and E. Zafiriou. Extensions and performance/robustness tradeoffs of the EWMA run-to-run controller by using the internal model control structure. *IEEE Transactions on Electronics Packaging Manufacturing*, 23:56–68, 2000.

[4] K.J. Astrom and B. Wittenmark. *Adaptive Control*. Addison-Wesley Publishing Company, Inc, 1995.

[5] J.S. Baras and N.S. Patel. A framework for robust run by run control with lot delayed measurements. *IEEE Transactions on semiconductor manufacturing*, 10(4):75–83, 1997.

[6] C.A. Bode. *Run-to-Run Control of Overlay and Linewidth in Semiconductor Manufacturing*. Ph.D. Dissertation, The University of Texas at Austin, 2001.

[7] D.S. Boning, W.P. Moyne, T.H. Smith, J. Moyne, R. Telfeyan, A. Hurwitz, S. Shellman, and J. Taylor. Run by run control of chemical-mechanical pol-

ishing. *IEEE Transactions on Semiconductor Manufacturing*, 19(4):307–314, 1996.

[8] G.E.P. Box and G.M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, Oakland, California, 1976.

[9] G.E.P. Box and T. Kramer. Statistical process monitoring and feedback adjustment – a discussion. *Technometrics*, 34:251–285, 1992.

[10] G.E.P. Box and A. Luceno. *Statistical process control by monitoring and feedback adjustment*. John Wiley and Sons, Inc., 1997.

[11] S.W. Butler and J.A. Stefani. Supervisory run-to-run control of a polysilicon gate etch using in situ ellipsometry. *IEEE Transactions on Semiconductor Manufacturing*, 7 (2):193–201, 1994.

[12] W. Campbell. *Model Predictive Run-to-Run Control of Chemical Mechanical Planarization*. Ph.D. Dissertation, The University of Texas at Austin, 1999.

[13] W.J. Campbell, S.K. Firth, A.J. Toprac, and T.F. Edgar. A comparison of run-to-run control algorithms. *Proceedings of the American Control Conference*, 3:2150–2155, 2002.

[14] E. Del Castillo. Long run and transient analysis of a doule EWMA feedback controller. *IIE Transactions*, 31:1157–1169, 1999.

[15] E. Del Castillo. *Statistical Process Adjustment for Quatlity Control*. John Wiley and Sons, 2002.

[16] E. Del Castillo and A. Hurwitz. Run-to-run process control: Literature review and extensions. *Journal of Quality Technology*, 29(2):184–196, 1997.

[17] E. Del Castillo and Ramkumar Rajagopal. A multivariate double EWMA process adjustment scheme for drifting processes. *IIE Transactions*, 34:1055–1068, 2002.

[18] E. Del Castillo and J.Y. Yeh. An adaptive run-to-run optimizing controller for linear and nonlinear semiconductor processes. *IEEE Transactions on Semiconductor Manufacturing*, 11 (2):285–295, 1998.

[19] A. Chen and R.S. Guo. Age-based double EWMA controller and its application to CMP processes. *IEEE Transactions on Semiconductor Manufacturing*, 14 (1):11–19, 2001.

[20] R. Chong, J. Wang, M.A. Purdy, D. Adams, P.G. Dye, and J.H. Hussey. Analysis of a run-to-run controller on a drifting STI etch process by augmentation of the integrated interferometric endpoint detection system. *Proceedings of the AEC/APC Conference*, 14, 2002.

[21] G. Chrystal. On the problem to construct the minimum circle enclosing n given points in the plane. *Proceedings of the Edinburgh Mathematical Society*, 3:30–33, 1885.

[22] A. Cohn. Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise. *Mathematische Zeitschrift*, 14:110–148, 1922.

[23] M. Denny. Watt steam governor stability. *European Journal of Physics*, 23(3):339–351, 2002.

[24] John C. Doyle and Gunter Stein. *IEEE Transactions on Automatic Control , NUMBER =*.

[25] T.F. Edgar, W.J. Campbell, and C. Bode. Model based control in microelectronics manufacturing. *Proceedings of the Conference on Decision and Control*, 38:4185–4191, 1999.

[26] J. Elzinga and D.W. Hearn. Geometric solutions for some minimax location problems. *Transportation Science*, 6:379–394, 1972.

[27] S.K. Firth, W.J. Campbell, and T.F. Edgar. Just-in-time adaptive disturbance estimation for run-to-run control of photolithography overlay. *Proceedings of SPIE Design, Process Integration, and Characterization for Microelectronics*, Santa Clara, 2002.

[28] R.L. Francis. Some aspects of a minimax location problem. *Operations Research*, 15(6):1163–1169, 1967.

[29] E. Frank. On the zeros of polynomials with complex coefficients. *Bulletin of the American Mathematical Society*, 52:144–157, 1946.

[30] M. Fujiwara. Über die algebraischen Gleichung deren wurzeln in einem Kreise oder in einer halbebene liegen. *Mathematische Zeitschrift*, 24:160–169, 1962.

[31] F.R. Gantmacher. *Applications of the Theory of Matrices*. Interscience Publishers, 1959.

[32] M.D. Greenberg. *Advanced Engineering Mathematics*. Prentice Hall, 1998.

[33] R.S. Guo, A. Chen, and J.J. Chen. Run-to-run control schemes for CMP process subject to deterministic drifts. *Semiconductor Manufacturing Technology Workshop*, pages 251–258, 2000.

[34] E.S. Hamby, P.T. Kabamba, and P.P. Khargonekar. A probabilistic approach to run-to-run control. *IEEE Transactions on Semiconductor Manufacturing*, 11(4):654–669, 1998.

[35] D.W. Hearn and J. Vijay. Efficient algorithms for the weighted minimum circle problem. *Operations Research*, 30 (4):777–795, 1982.

[36] A. Hurwitz. Über die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt. *Mathematicshe Annalen*, 46:273–284, 1895.

[37] A. Ingolfsson and E. Sachs. Stability and sensitivity of an EWMA controller. *Journal of Quality Technology*, 25(4):271–287, 1993.

[38] M. Janakiram and J.B. Keats. Combining SPC and EPC in a hybrid industry. *Journal of Quality Technology*, 30:189–200, 1998.

[39] P. Jowett and Victor Morozov. Shallow trench isolation run-to-run control project at Infenion Technologies Richmond. *IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, pages 107–112, 2002.

[40] E.I. Jury. A simplified stability criterion for linear discrete systems. *IRE Proceedings*, 50(6):1493–1500, 1962.

[41] K.Chamness, G. Cherry, R. Good, and S.J. Qin. A comparison of run to run control algorithms for the CMP with metrology delays. *Proceedings of the AEC/APC Symposium*, 13, 2001.

[42] J. B. Keats and D. C. Montgomery. *Statistical Applications in process control*. Marcel Dekker, Inc., 1996.

[43] R.L. Kosut, D. de Roover, A. Emani-Naeini, and J.L. Ebert. Run-to-run control of static systems. *Proceedings of the IEEE Conference on Decision and Control*, 37:695–700, 1998.

[44] J.R. Leigh. *Applied Digital Control*. Prentice Hall International, 1984.

[45] F.L. Lewis and V.L. Syrmos. *Optimal Control*. John Wiley and Sons, 1995, 2nd ed.

[46] J.C. Maxwell. On governors. *Proceedings of the Royal Society London*, 16:270–283, 1868.

[47] M.L. Miller. Impact of multi-product and process manufacturing on run-to-run control. In A. Ghanbari and A. Toprac, editors, *Process, Equipment, and Materials Control In Integrated Circuits Manufacturing, III*. SPIE, 1997.

[48] Z. Ning, J.R. Moyne, T. Smith, D. Boning, E. Del Castillo, J.Y. Yeh, and A. Hurwitz. A comparative analysis of run-to-run control algorithms in the semiconductor manufacturing industry. *Proceedings of the Advanced Semiconductor Manufacturing Conference*, pages 375–381, 1996.

[49] B.J. Oommen. An efficient geometric solution to the minimum spanning cirlce problem. *Operations Research*, 35(1):80–86, 1987.

[50] Z.J. Palmor. Properties of optimal stochastic control systems with dead-time. *Automatica*, 18(1):107–116, 1982.

[51] A.J. Pasadyn. *Simultaneous control and identification for multiple product and process environments in semiconductor manufacturing.* Ph.D. Dissertation, The University of Texas at Austin, 2001.

[52] S.J. Qin, G.W. Scheid, and T.J. Riley. Adaptive run to run control for intermittent batch operations. *Proceedings of the American Control Conference*, 3:2168–2173, 2002.

[53] E.J. Routh. *Dynamics of a System of Rigid Bodies.* Macmillan, 1892.

[54] E. Sachs, R.S. Guo, S. Ha, and A. Hu. Process control system for VLSI fabrication. *IEEE Transactions on Semiconductor manufacturing*, 4(2):134–143, 1991.

[55] E. Sachs, A. Hu, and A. Ingolfsson. Run by run process control: Combining SPC and feedback control. *IEEE Transactions on Semiconductor Manufacturing*, 8:26–43, 1995.

[56] I. Schur. Über Potenzreihen die im Innern des Einheitskreises beschränkt sind. *Journal für die Reine und Angewandte Mathematik*, 147:205–232, 1917.

[57] D. E. Seborg, T. F. Edgar, and D. A. Mellichamp. *Process Dynamics and Control.* John Wiley & Sons, 1989.

[58] F.G. Shi and B. Zhao. Fundamentals of CMP for IC manufacturing: recent progress. *Proceedings of the Solid State and Integrated Circuit Technology Conference*, 5:113–115, 1998.

[59] T. Smith, D. Boning, J. Stefani, and S. Butler. Run by run advanced process control of metal sputter deposition. *IEEE Transactions of Semiconductor Manufacturing*, 11(2):276–284, 1998.

[60] T. Smith and D.Boning. A self-tuning EWMA controller utilizing artificial neural network function approximation techniques. *IEEE Transactions of Components, Packaging, and Manufacturing Technology*, 20(2):121–132, 1997.

[61] T.H. Smith. *Device Independent Process Control of dielectric chemical mechanical planarization.* Ph.D. Dissertation, The Massachussetts Institute of Technology, 1999.

[62] J.J. Sylvester. A question in the geometry of situation. *Quarterly Journal of Pure and Applied Mathematics*, 1:79, 1857.

[63] S.T. Tseng, R.J. Chou, and S.P. Lee. A study on a multivariate EWMA controller. *IIE Transactions*, 34:541–549, 2002.

[64] S.T. Tseng, A.B. Yeh, F.T. Tsung, and Y.Y. Chan. A study of a variable EWMA controller. *IEEE Transactions on Semiconductor Manufacturing*, 16(4):633–643, 2003.

[65] W.T. Tucker, F.W. Faltin, and S.A. Vander Wiel. Algorithmic statistical process control: concepts and application. *Technometrics*, 35:278–281, 1993.

[66] J. Wang and S.J. Qin. EWMA, Kalman filter and recursive least squares - relationships and modifications. *The Texas-Wisconsin Modelling and Control Consortium Spring Meeting, Austin Texas*, 2003.

[67] S.A. Vander Weil, W.T. Tucker, F.W. Faltin, and N. Doganaksoy. Algorithmic statistical process control: Concepts and an application. *Technometrics*, 34:286–297, 1992.

[68] Y. Wu, J. Gilhooly, and B. Philips. Control methods for the chemical mechanical polishing process in shallow trench isolation. *Advanced Semiconductor Manufacturing Conference and Workshop*, pages 66–70, 1998.

[69] E. Zafiriou, R.A. Adomaitis, and G. Gattu. An approach to run-to-run control for rapid thermal processing. *Proceedings of the American Control Conference*, 2:1286–1288, 1995.

[70] C. Zhang, H. Deng, and J.S. Baras. Performance evaluation of run-to-run control methods in semiconductor processes. *Proceedings of the 2003 IEEE Conference on Control Applications*, 1:841–848, 2003.

Vita

Richard Paul Good was born in Atlanta, Georgia on 25 August 1977, the son of Edwin L. Good and Esther B. Good. He received a Bachelor of Science degree in Chemical Engineering from the Georgia Institute of Technology in the spring of 2000 and, after spending a summer at Merck and Co., he entered graduate school at the University of Texas at Austin in the fall of 2000. Richard received a Master of Science degree in Chemical Engineering from the University of Texas at Austin in 2002 and began working at Advanced Micro Devices shortly thereafter in the Advanced Process Control group. Richard currently resides in Dresden, Germany, with his wife Kara, where he is helping to develop the advanced process control framework for AMD's Fab36.

Permanent address: Boltenhagener Straße 20
Dresden, Germany 01109

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.